Central Core Laboratory versus Site Interpretation of Coronary CT Angiography: Agreement and Association with Cardiovascular Events in the PROMISE Trial[1]

Michael T. Lu, MD, MPH
Nandini M. Meyersohn, MD
Thomas Mayrhofer, PhD
Daniel O. Bittner, MD
Hamed Emami, MD
Stefan B. Puchner, MD
Borek Foldyna, MD
Martin E. Mueller
Steven Hearne, MD
Clifford Yang, MD
Stephan Achenbach, MD
Quynh A. Truong, MD, MPH
Brian B. Ghoshhajra, MD, MBA
Manesh R. Patel, MD
Maros Ferencik, MD, PhD
Pamela S. Douglas, MD
Udo Hoffmann, MD, MPH

Purpose: To assess concordance and relative prognostic utility between central core laboratory and local site interpretation for significant coronary artery disease (CAD) and cardiovascular events.

Materials and Methods: In the Prospective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) trial, readers at 193 North American sites interpreted coronary computed tomographic (CT) angiography as part of the clinical evaluation of stable chest pain. Readers at a central core laboratory also interpreted CT angiography blinded to clinical data, site interpretation, and outcomes. Significant CAD was defined as stenosis greater than or equal to 50%; cardiovascular events were defined as a composite of cardiovascular death or myocardial infarction.

Results: In 4347 patients (51.8% women; mean age ± standard deviation, 60.4 years ± 8.2), core laboratory and site interpretations were discordant in 16% (683 of 4347), most commonly because of a finding of significant CAD by site but not by core laboratory interpretation (80%, 544 of 683). Overall, core laboratory interpretation resulted in 41% fewer patients being reported as having significant CAD (14%, 595 of 4347 vs 23%, 1000 of 4347; $P < .001$). Over a median follow-up period of 25 months, 1.3% (57 of 4347) sustained myocardial infarction or cardiovascular death. The *C* statistic for future myocardial infarction or cardiovascular death was 0.61 (95% confidence interval [CI]: 0.54, 0.68) for the core laboratory and 0.63 (95% CI: 0.56, 0.70) for the sites.

Conclusion: Compared with interpretation by readers at 193 North American sites, standardized core laboratory interpretation classified 41% fewer patients as having significant CAD.

Online supplemental material is available for this article.

Clinical trial registration no. NCT01174550.

[1] From the Cardiac PET MR CT Program, Massachusetts General Hosp and Harvard Medical School, Boston, Mass (M.T.L., N.M.M., T.M., D.O.B., H.E., S.B.P., B.B.G., B.F., M.E.M., M.F., U.H.); School of Business Studies, Stralsund Univ of Applied Sciences, Stralsund, Germany (T.M.); Dept of Internal Medicine (Cardiology), Friedrich Alexander Univ Hosp, Erlangen, Germany (D.O.B., S.A.); Dept of Angiography and Interventional Radiology, Medical Univ Vienna, Vienna, Austria (S.B.P.); Delmarva Health LLC, Salisbury, Md (S.H.); Dept of Radiology, Univ of Connecticut Health Ctr, Farmington, Conn (C.Y.); Dalio Institute of Cardiovascular Imaging, Weill Cornell Medical College, New York, NY (Q.A.T.); Duke Clinical Research Inst, Duke Univ School of Medicine, Durham, NC (M.R.P., P.S.D.); and Knight Cardiovascular Inst, Oregon Health & Science Univ, Portland, Ore (M.F.). From the 2017 RSNA Annual Meeting. Received Aug 12, 2017; revision requested Oct 16; revision received Oct 31; accepted Oct 31; final version accepted Oct 31. **Address correspondence to** M.T.L. (e-mail: *mlu@mgh.harvard.edu*).

The parent PROMISE trial was supported by National Heart, Lung, and Blood Institute grants (R01HL098237, R01HL098236, R01HL098305, and R01HL098235). The funding sources had no role in the design of the study, data analysis, or decision to submit the manuscript. The content of this manuscript is solely the responsibility of the authors and does not necessarily reflect the views of the National Heart, Lung, and Blood Institute.

M.T.L. and N.M.M. contributed equally to this work.

Detection of significant coronary artery disease (CAD) at coronary computed tomographic (CT) angiography yields important prognostic information and may trigger increased downstream testing or revascularization (1). As a pragmatic comparative effectiveness trial, the Prospective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) trial relied on local CT readers at 193 North American sites to interpret coronary CT angiography for significant CAD (2). Local interpretation of coronary CT angiography reflects actual clinical practice, but may vary in terms of reader expertise, risk averseness, and willingness to read through calcification and other artifacts (3). In addition, access to clinical risk factors and the results of other diagnostic tests may introduce test interpretation bias (4). These interpretive challenges are not unique to coronary CT angiography and are common to all imaging tests. In contrast, central core laboratory analysis provides a standardized assessment blinded to clinical information, which may enable a less-biased view of the actual extent of CAD. Although it is challenging to define a standard of reference between site and core laboratory interpretations, PROMISE provides the opportunity to make a comparison based on prognostic ability to predict future myocardial infarction and cardiovascular death. Reports on differences between site and core laboratory interpretation of diagnostic imaging are scarce, with only a few invasive coronary angiography (ICA) studies relating differences to clinical outcomes (5–7). The purpose of our study was to assess concordance and relative prognostic utility of central core laboratory versus local site interpretation for significant CAD. The primary hypothesis of our study was that the core laboratory readers would find less significant CAD, defined as stenosis greater than or equal to 50%, than would local site readers. The secondary hypothesis was that there would not be a significant difference in discriminatory power for myocardial infarction and cardiovascular death based on these interpretive differences.

Advances in Knowledge

- Core laboratory interpretation of coronary CT angiography classified 41% fewer patients as having significant coronary artery disease (CAD) compared with interpretation by readers at 193 North American sites.
- The *C* statistic for future myocardial infarction or cardiovascular death was 0.61 for the core laboratory and 0.63 for the sites.
- Discordance between core laboratory and site interpretation increased with coronary artery calcium score, but not with body mass index or heart rate.
- Both level II (experience reading ≥150 coronary CT angiograms) and level III (≥300 coronary CT angiograms) site readers more frequently found significant CAD than did the core laboratory readers.

Materials and Methods

Study Design and Population

In this prespecified retrospective analysis, we assessed concordance between site and core laboratory interpretations of coronary CT angiography for significant CAD and their association with cardiovascular outcomes. The analysis was embedded in PROMISE, a pragmatic comparative effectiveness trial of coronary CT angiography versus functional testing strategies conducted between July 2010 and September 2013 (2). In PROMISE, 10003 stable symptomatic outpatients without known CAD were randomized to anatomic (coronary CT angiography) or functional (nuclear stress or stress echocardiography) testing arms at 193 North American sites representing a spectrum of community and academic practices (2,3). Participants provided written informed consent; local and central institutional review boards approved this Health Insurance Portability and Accountability Act–compliant study.

The current analysis included participants enrolled in PROMISE who were randomized to the CT angiography arm and underwent CT angiography as their initial diagnostic test. Participants who did not undergo testing, who underwent another diagnostic test before CT angiography, who underwent a noncontrast material–enhanced CT only for coronary artery calcium (CAC) scoring (no CT angiography), whose site interpretation of CT angiography was unavailable, whose CT angiograms were unavailable, or whose CT angiogram data sets were deemed of nondiagnostic quality by either the site reader or the core laboratory reader were excluded (Fig 1).

Implication for Patient Care

- The data suggest an opportunity to report fewer positive findings at coronary CT angiography, with the potential to decrease downstream testing.

https://doi.org/10.1148/radiol.2017172181

Content code: CA

Radiology 2018; 287:87–95

Abbreviations:

CAC = coronary artery calcium
CAD = coronary artery disease
CI = confidence interval
ICA = invasive coronary angiography
PROMISE = Prospective Multicenter Imaging Study for Evaluation of Chest Pain
QCA = quantitative coronary angiography

Author contributions:

Guarantors of integrity of entire study, M.T.L., P.S.D., U.H.; 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, M.T.L., N.M.M., D.O.B., S.B.P., B.F., B.B.G., M.R.P., M.F., P.S.D., U.H.; clinical studies, M.T.L., N.M.M., D.O.B., S.B.P., S.H., C.Y., Q.A.T., B.B.G., M.R.P., M.F., P.S.D., U.H.; experimental studies, M.R.P., M.F.; statistical analysis, N.M.M., T.M., B.F., M.F., P.S.D.; and manuscript editing, M.T.L., N.M.M., D.O.B., S.B.P., B.F., M.E.M., S.H., C.Y., S.A., Q.A.T., B.B.G., M.R.P., M.F., P.S.D., U.H.

Conflicts of interest are listed at the end of this article.

Figure 1

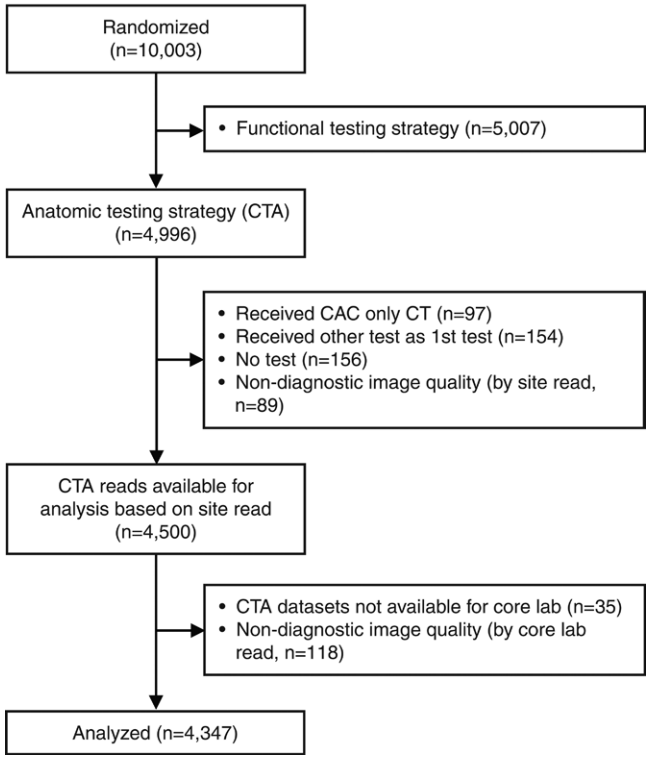

Figure 1: Flowchart of patient enrollment from the anatomic coronary CT angiography (CTA) arm of the PROMISE trial.

Previous PROMISE reports compared the outcomes of anatomic CT angiography versus functional testing strategies for stable chest pain and the prognostic value of anatomic CT angiography versus functional test findings (1,2). Both prior studies relied on site interpretation. In the current article, we report differences between core laboratory and site interpretation of CT angiography and their discriminatory power to predict myocardial infarction and cardiovascular death.

Coronary CT Angiography

Coronary CT angiography was performed as part of the clinical evaluation of chest pain by using CT scanners with at least 64 detector rows according to the guidelines of the Society of Cardiovascular Computed Tomography (8). CT imaging for CAC scoring was encouraged but not required as part of the CT angiography protocol. Five vendors (General Electric, Boston, Mass; Hitachi, Twinsburg, Ohio; Phillips, Amsterdam, the Netherlands; Siemens, Erlangen, Germany; and Toshiba, Otawara, Japan) and multiple generations of scanners were represented.

Site Interpretation of CT Angiography

Coronary CT angiography was interpreted locally in the context of the clinical workup for CAD. Contemporaneous clinical information, including cardiovascular risk factors and the results of other testing, was available to the interpreting physician; clinical management decisions were made on the basis of site interpretation of CT angiography. Site readers were certified by the CT core laboratory prior to activation in the trial (3). As part of this certification process, 582 site readers self-reported their coronary CT angiography expertise as either level II or level III based on the American College of Cardiology, American Heart Association, and Society of Cardiovascular Computed Tomography criteria, with experience interpreting greater than or equal to 150 CT angiograms for level II and greater than or equal to 300 CT angiograms for level III (9,10). Site interpretations were prospectively collected, with significant CAD defined as a qualitative diameter stenosis greater than or equal to 50% per patient.

Central Core Laboratory Interpretation of CT Angiography

At the conclusion of the trial, coronary CT angiograms were transferred to the CT core laboratory for analysis. Coronary CT angiograms were randomly assigned to one of six level III core laboratory readers (M.T.L., N.M.M., D.O.B., H.E., S.B.P., and M.F., with 3–10 years of experience interpreting coronary CT angiography). By using dedicated three-dimensional coronary analysis software (AQi, version 4.4.8; TeraRecon, Foster City, Calif) and following the interpretive guidelines of the Society of Cardiovascular Computed Tomography, core laboratory readers first qualitatively graded CT angiography image quality as either diagnostic or nondiagnostic (11). All CT angiograms were then read qualitatively for significant CAD, defined as at least one diameter stenosis greater than or equal to 50% per patient. Core laboratory readers also analyzed 50 randomly selected CT angiography data sets to determine interobserver variability for significant CAD (κ of 0.69), similar to another multicenter trial of coronary CT angiography (12). Core laboratory readers were blinded to all aspects of clinical care including risk factors, the site interpretation, the results of other diagnostic testing, interventions, and clinical outcomes. The results of the core laboratory interpretation were not available to care providers and did not play a role in clinical management.

Cardiovascular Events

The primary clinical end point was defined as a composite of myocardial infarction or cardiovascular death.

Patients returned to sites at 60 days, and were subsequently contacted via phone or mail at 6-month intervals after randomization for a minimum of 1 year and up to 4 years (median, 25 months). A blinded, independent clinical events committee adjudicated all potential cardiovascular events based on prospectively defined guidelines (2,3).

ICA and Quantitative Coronary Angiography

Site physicians independently decided whether to refer patients to ICA based on the results of site interpretation of CT angiography, other diagnostic testing, and clinical risk factors. At the conclusion of the trial, ICA images were transferred to a second ICA core laboratory wholly separate from the CT angiography core laboratory. The ICA core laboratory readers performed quantitative coronary angiography (QCA) diameter stenosis measurements blinded to clinical information and the results of all other testing by using dedicated software (QAngio XA, version 7.3; MEDIS Medical Imaging Systems, Leiden, the Netherlands) as previously described (5). QCA results were not available to care providers and did not affect clinical management.

Statistical Analysis

Continuous variables are presented with means and standard deviations or medians and interquartile ranges; comparisons between groups for nonpaired data were performed with the independent sample Student *t* test or Wilcoxon rank-sum test. Categorical variables are presented as frequencies and percentages; comparisons for nonpaired data were made by using the Fisher exact test. The primary analysis assessed concordance between core laboratory and site interpretation of coronary CT angiography for significant CAD. Contingency tables and agreement rates were presented for core laboratory and site interpretation for per-patient significant CAD (≥50% luminal stenosis) and compared by using the McNemar test. The secondary analysis compared discrimination for cardiovascular events by using the log-rank test, first by concordance category (concordant for no significant CAD, concordant for significant CAD, discordant with significant CAD by core laboratory interpretation but not site interpretation, and discordant with significant CAD by site interpretation but not core laboratory interpretation) and then by comparing site versus core laboratory findings of significant CAD. Kaplan-Meier curves were generated for each concordance category and compared by using the log-rank test. The Harrell *C* concordance statistic (*C* statistic) was calculated based on Cox proportional hazard regression models by using significant CAD (binary variable) as an independent variable (13,14).

A sensitivity analysis based on a severe CAD (≥50% left main or ≥70% other coronary artery stenosis) threshold was performed. Subgroup analyses were performed based on site reader expertise (site level II and level III vs core laboratory) and patient factors potentially affecting the interpretation of coronary CT angiography, including CAC score by site read (CAC score: 1–99, 100–399, and ≥400), body mass index (kilograms per square meter: <25, 25–29, 30–34, ≥35), and heart rate during CT angiography acquisition (beats per minute: <60, 60–69, 70–79, and ≥80). A second subanalysis in the subgroup of patients undergoing QCA assessed the accuracy of core laboratory versus site interpretation of CT angiography by using core laboratory QCA greater than or equal to 50% stenosis at ICA as the standard of reference. For each stenosis category, the sensitivity, specificity, positive predictive value, and negative predictive value, including 95% confidence intervals (CIs) by using the exact binominal distribution, were calculated. The McNemar test was used to compare sensitivity and specificities, whereas the method by DeLong was used to compare areas under the receiver operating characteristics curves (15). Interobserver variability between the six core laboratory readers was calculated by using the κ statistic. A two-sided *P* value of ≤ .05 was considered to indicate statistical significance. Statistical analysis was performed by using Stata (version 14.2; StataCorp, College Station, Tex).

Results

Study Population

The analysis included 4347 patients (51.8% women; mean age ± standard deviation, 60.4 years ± 8.2). Baseline characteristics stratified by core laboratory and site interpretation are described in Table E1 (online).

Core Laboratory and Site Concordance

Contingency tables describing concordance between core laboratory and site interpretations for binary significant CAD (≥50% stenosis) are provided in Table 1. An additional contingency table including categories of 0%, 1%–49%, and greater than or equal to 50% stenosis is provided in Table E2 (online). Discordant interpretations between site and core laboratory readers were more likely in patients with a higher cardiovascular risk profile, with higher median Framingham risk scores compared with those with concordant interpretations (22.5 [interquartile range, 14.4–33.7] vs 16.0 [interquartile range, 9.9–26.2]; $P < .001$).

Core laboratory and site interpretations were discordant in 16% (683 of 4347) and concordant in 84% (3664 of 4347). Eighty percent (544 of 863) of discordant interpretations were in patients who had significant CAD by site but not by core laboratory interpretation. The core laboratory readers found significant CAD in 41% fewer patients compared with site readers (absolute difference, −9%; 14% [595 of 4347] vs 23% [1000 of 4347]; $P < .001$).

Cardiovascular Events and Association with Concordance

Over a median follow-up period of 25 months, 1.3% (57 of 4347) sustained a cardiovascular event (myocardial infarction or cardiovascular death). There was a graded increase in event rate based on core laboratory versus site concordance: 0.8% (27 of 3208) in patients with concordance for no significant CAD, 2% (14 of 683) in

Table 1

Core Laboratory and Local Site Interpretation of Coronary CT Angiography

Parameter	Core Laboratory <50% Stenosis	Core Laboratory ≥50% Stenosis	Total
Agreement for interpretation of significant CAD			
Site <50% stenosis	74 (3208/4347)	3 (139/4347)	77 (3347/4347)
Site ≥50% stenosis	13 (544/4347)	10 (456/4347)	23 (1000/4347)
Total	86 (3752/4347)	14 (595/4347)	…
Cardiovascular events (cardiovascular death or myocardial infarction)			
Site <50% stenosis	0.8 (27/3208)	2.2 (3/139)	0.9 (30/3347)
Site ≥50% stenosis	2.0 (11/544)	3.5 (16/456)	2.7 (27/1000)
Total	1.0 (38/3752)	3.2 (19/595)	…

Note.—Significant CAD was defined as stenosis greater than or equal to 50%. Data are percentages, with numerators and denominators in parentheses.

Figure 2

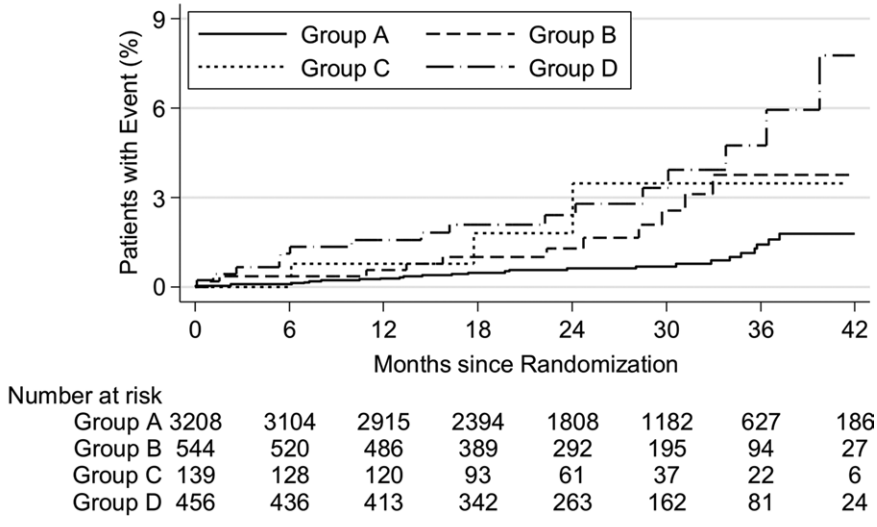

Figure 2: Graph shows Kaplan-Meier curves for cardiovascular events (cardiovascular death or myocardial infarction) based on concordance of local site and core laboratory interpretation of coronary CT angiography. Group A = concordant for no significant CAD; Group B = discordant, with significant CAD by site interpretation but not by core laboratory interpretation; Group C = discordant, with significant CAD by core laboratory interpretation but not site interpretation; and Group D = concordant for significant CAD.

patients with discordance, and 3.5% in patients (16 of 456) with concordance for significant CAD ($P < .001$) (Table 1). For patients with discordant interpretations, the event rate was 2.0% (11 of 544) when the core laboratory readers found no significant CAD but the site readers found significant CAD, and 2.2% (three of 139) when the core laboratory readers found significant CAD but the site readers found no significant CAD ($P = .76$). Kaplan-Meier curves for cardiovascular events in the four categories of agreement are provided in Figure 2. The *C* statistic for future myocardial infarction or cardiovascular death was 0.61 (95% CI: 0.54, 0.68) for the core laboratory and 0.63 (95% CI: 0.56, 0.70) for the sites.

Sensitivity Analysis Based on Severe CAD Threshold

A sensitivity analysis by using a severe CAD threshold (≥50% left main or ≥70% other coronary artery) found similar differences between site and core laboratory interpretations (see Appendix E1 [online]).

Subgroup Analysis Based on Site Reader Expertise

Site reader information was available for 2430 CT angiograms, with 1025 CT angiograms read by level II readers and 1405 by level III readers (Table 2). Level II and III site readers had similar overall concordance rates with the core laboratory readers (82%, 839 of 1025 and 84%, 1179 of 1405, respectively; $P = .19$). Level II site readers found more significant CAD than did level III site readers (33%, 337 of 1025 vs 26%, 363 of 1405; $P < .001$). Both level II (33%, 337 of 1025 vs 19%, 193 of 1025; $P < .001$) and level III (26%, 363 of 1405 vs 15%, 217 of 1405; $P < .001$) site readers called significant CAD more frequently than did the core laboratory readers.

As seen in the overall group, for both level II and III readers there was a graded increase in events based on concordance with the core laboratory and significant CAD category (Table 2). For CT images read by level II readers, the *C* statistic for future myocardial infarction or cardiovascular death was 0.68 (95% CI: 0.54, 0.81) for the core laboratory and 0.65 (95% CI: 0.52, 0.78) for the sites. For CT images read by level III readers, the *C* statistic for future myocardial infarction or cardiovascular death was 0.61 (95% CI: 0.51, 0.70) for the core laboratory and 0.66 (95% CI: 0.57, 0.76) for the sites.

Patient Factors Potentially Affecting Concordance

We focused on patient factors known to be associated with diagnostic accuracy,

Table 2

Subanalysis Stratified by Site Reader Expertise (Level II or Level III)

Parameter	Core Laboratory <50% Stenosis	Core Laboratory ≥50% Stenosis	Total
Agreement for interpretation of significant CAD			
Level II readers*			
Site <50% stenosis	65 (667/1025)	2 (21/1025)	67 (688/1025)
Site ≥50% stenosis	16 (165/1025)	17 (172/1025)	33 (337/1025)
Total	81 (832/1025)	19 (193/1025)	
Level III readers*			
Site <50% stenosis	71 (1002/1405)	3 (40/1405)	74 (1042/1405)
Site ≥50% stenosis	13 (186/1405)	13 (177/1405)	26 (363/1405)
Total	85 (1188/1405)	15 (217/1405)	…
Cardiovascular events (myocardial infarction or cardiovascular death)			
Level II readers*			
Site <50% stenosis	0.9 (6/667)	9.5 (2/21)	1.2 (8/688)
Site ≥50% stenosis	1.8 (3/165)	2.9 (5/172)	2.4 (8/337)
Total	1.1 (9/832)	3.6 (7/193)	
Level III readers*			
Site <50% stenosis	1.2 (12/1002)	0.0 (0/40)	1.2 (12/1042)
Site ≥50% stenosis	3.2 (6/186)	5.7 (10/177)	4.4 (16/363)
Total	1.5 (18/1188)	4.6 (10/217)	…

Note.—Significant CAD was defined as stenosis greater than or equal to 50%. Data are percentages, with numerators and denominators in parentheses.

* There were 1025 level II readers and 1405 level III readers.

including extent of CAC, heart rate, and body mass index. CAC was present in 2497 patients. Discordance between site and core laboratory interpretations increased with increasing CAC category ($P < .001$) (Fig 3, Table 3). Neither body mass index nor heart rate during CT angiography were associated with discordant CT angriogram interpretations (Fig E1 [online]).

Accuracy with Invasive Quantitative Coronary Angiography as the Reference Standard

ICA was performed in 491 patients, with 458 undergoing QCA interpretation. In these 458 patients, QCA stenosis greater than or equal to 50% was present in 51% (233 of 458). With QCA stenosis greater than or equal to 50% as the reference standard, core laboratory interpretation of CT angiography had lower sensitivity (73% [95% CI: 67%, 79%] vs 98% [95% CI: 96%, 100%]; $P < .001$), lower negative predictive value (70% [95% CI: 65%, 74%] vs 90% [95% CI: 76%, 97%]; $P = .010$), higher specificity (65% [95% CI: 58%, 71%] vs 16% [95% CI: 12%, 22%]; $P < .001$), and higher positive predictive value (68% [95% CI: 62%, 74%] vs 55% [95% CI: 50%, 60%]; $P = .001$) than did sites. As a measure of overall accuracy, the area under the receiver operating characteristics curves was greater for the core laboratory interpretation of CT angiography than for sites (0.69 vs 0.57; $P < .001$).

Discussion

The goal of our study was to assess agreement between standardized central core laboratory and site interpretations of coronary CT angiography and their relative prognostic utility. Agreement for and against significant CAD was associated with high and low rates of cardiovascular events, respectively. Disagreement was associated with an intermediate rate of events. Core laboratory interpretation classified 41% fewer patients as having significant CAD compared with site interpretation.

Clinical trials of diagnostic imaging focus on comparing new versus established technology. Our results underscore the importance of also considering who interprets these tests and how they are read. PROMISE was a pragmatic trial that compared anatomic CT angiography and functional stress testing strategies; in this context, relying on local site physicians to interpret CT angiography is appropriate and reflects actual practice. With local readers from 193 North American sites, PROMISE constitutes a generalizable representation of CT angiography reader reporting practice and an opportunity to assess the strengths and weaknesses of CT angiography interpretation.

The major weakness of CT angiography is a relatively low positive predictive value, which leads to a high rate of referral to ICA compared with functional testing (16,17). In PROMISE, patients randomized to coronary CT angiography were referred to ICA 51% more often than were the functional arm (12.2% vs 8.1%), yet there was no improvement in 2-year cardiovascular outcomes (2). In this context, the most important result of this analysis is that core laboratory interpretation classified 41% fewer patients as having significant CAD. The data suggest the potential for greater education and expertise to reduce positive interpretations and downstream testing.

A strength of our study was that the comparison of core laboratory and site interpretations can be extended to a comparison of their prognostic value for independently adjudicated "hard" cardiovascular events (myocardial infarction or cardiovascular death). Myocardial infarction and cardiovascular death are generally considered unbiased by the results of diagnostic testing, as opposed to "soft" outcomes, such as hospitalization for unstable angina and revascularization, that are influenced by the site interpretation of CT angiography. Significant CAD at CT angiography

Figure 3

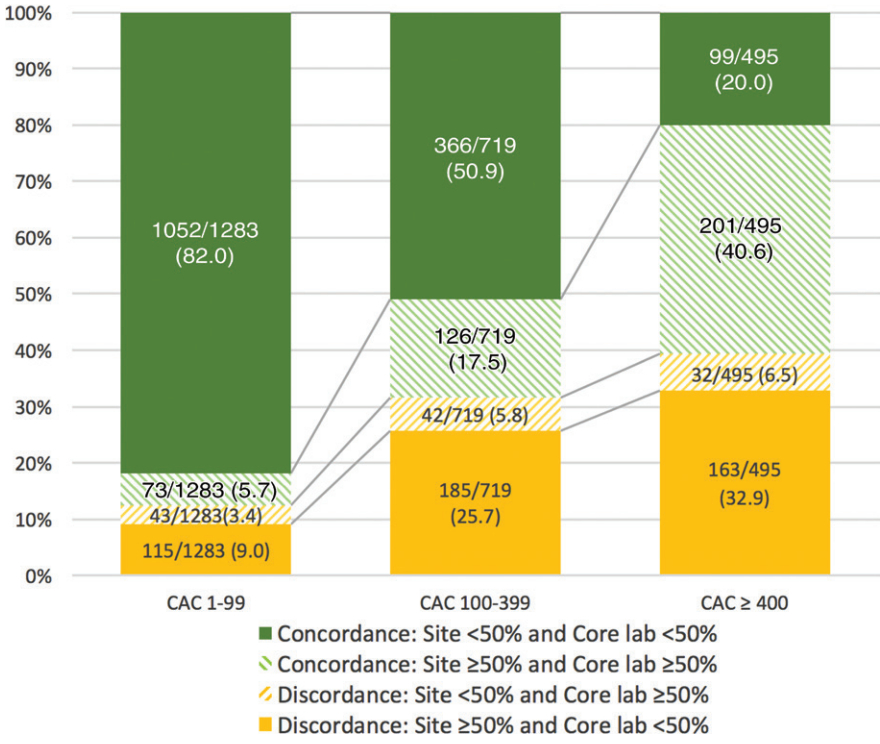

Figure 3: Bar graph shows core laboratory and site agreement for significant CAD (≥50% stenosis) by CAC category. Greater CAC was associated with greater rates of discordant interpretations ($P < .001$). Values in parentheses are percentages.

Table 3

Site and Core Laboratory Interpretation of Significant CAD by CAC Category

Parameter	CAC 1–99	CAC 100–399	CAC ≥400	Total
Site	14.7 (188/1283)	43.3 (311/719)	73.5 (364/495)	34.6 (863/2497)
Core laboratory	9.0 (116/1283)	23.4 (168/719)	47.1 (233/495)	20.7 (517/2497)

Note.—Significant CAD was defined as stenosis greater than or equal to 50%. Data are percentages, with numerators and denominators in parentheses.

was a predictor of myocardial infarction and cardiovascular death for both core laboratory and site interpretation (*C* statistics: core laboratory, 0.61 [95% CI: 0.54, 0.68] and site, 0.63 [95% CI: 0.56, 0.70]). This was a secondary retrospective analysis of the PROMISE trial, and given the small number of events (1.3%, 54 of 4347), it was not powered to conclude whether core laboratory and site interpretation had different discriminatory value for myocardial infarction and cardiovascular death.

Potential systematic differences between core laboratory and site interpretations deserve discussion. First, there were differences in reader experience. Core laboratory readers had level III expertise, whereas at sites 42% of CT angiograms were interpreted by less experienced level II readers. In the subanalysis stratified by site reader expertise, level II site readers called significant CAD in one-third of patients, compared with one-quarter for level III site readers. More experienced CT angiography readers are more accurate than are less experienced readers with ICA as the standard of reference (18–20). With greater experience also comes greater confidence in calling borderline cases negative, and this may in part explain why less experienced readers called more significant CAD. To qualify as level III, readers had to have interpreted at least 300 CT angiograms as per the American College of Cardiology, American Heart Association, and Society of Cardiovascular Computed Tomography guidelines, which are in line with similar criteria of 300 examinations for the highest level of expertise in echocardiography and nuclear cardiology (21,22). However, 300 CT angiograms is a relatively low ceiling unlikely to reflect the full spectrum of expertise (18,20). This threshold is similar to the American College of Radiology Cardiac CT Certificate of Advanced Proficiency (500 cardiac CT examinations and proficiency test) and the Certifying Examination in Cardiovascular Computed Tomography (150 cardiac CT examinations and proficiency test) (10,23).

Yet in CT angiograms read by level III site readers, the core laboratory interpretation still found 42% less significant CAD than did site interpretation (26% vs 15%). This suggests additional factors beyond expertise. Access to clinical information (test interpretation bias) may contribute (4). Interpretation of clinical coronary CT angiography occurs in the context of patient treatment, with access to patient risk factors and the results of other diagnostic tests. That incorporating pretest probability aids the interpretation of diagnostic testing is foundational in medicine and radiology. However, when risk scores systematically overestimate actual risk, as was the case in PROMISE and many modern cardiovascular trials (24,25), access to clinical information may lead to overestimation of the extent of CAD. Third, clinical readers may read coronary CT angiography only a few days per month, and then only a few per day. Cases are infrequently read without interruption. In contrast, the core laboratory is an artificial environment designed to maximize diagnostic accuracy

and limit bias. Cases are consecutively read in batches and without interruption. Fourth, current US practice incentives punish false-negative findings (ie, lawsuits) while accepting a high rate of false-positive findings. There is little reward for parsimony. These factors are not unique to coronary CT angiography and may also bias interpretation of other imaging tests.

We further assessed the effect of patient factors on discordance. Not surprisingly, discordant interpretations were more common with higher CAC scores. Several factors likely contribute. First, CAC is a surrogate for CAD, and the more CAD is present, the greater the opportunity for discrepant interpretations. Second, calcium blooming artifact makes it more difficult to evaluate the coronary lumen (26), and a high CAC score is a strong predictor of future cardiovascular events (27). Core laboratory readers are less risk averse and may be more willing to read through calcium. Other patient factors known to affect image quality, including heart rate and body mass index, did not affect discordance.

In the 458 patients who went on to undergo ICA, we compared the diagnostic accuracy of qualitative core laboratory and site interpretation of CT angiography with core laboratory QCA stenosis greater than or equal to 50% as the reference standard. Core laboratory interpretation of CT angiography had worse sensitivity but better specificity than did sites, as would be expected given the lower number of positive interpretations in the entire population. As a measure of accuracy, the core laboratory interpretation of CT angiography had a greater area under the receiver operating characteristics curves than did the sites (0.69 vs 0.57, $P < .001$). It should be emphasized that less than half of patients with significant CAD by site interpretation of CT angiography were referred to ICA. PROMISE was a pragmatic trial, and after randomization to CT angiography, site physicians were free to manage their patients based on all of the available clinical data. Presumably, the decision to refer patients to ICA was based not only on the CT angiography results, but also on clinical risk factors, the results of other diagnostic tests, and patient and physician preference. This potential for bias should be considered when critically evaluating this accuracy comparison. Also, in a PROMISE analysis including 929 patients undergoing ICA from both the CT angiography and functional testing arms, there were substantial discrepancies between core laboratory QCA and visual estimation of stenosis at ICA by sites, with a 19% disagreement rate for significant CAD. Similar to our findings with coronary CT angiography, at ICA the core laboratory readers found 28% fewer patients with significant CAD than did site readers; event rates were highest (5.1%) when the site and core laboratory readers agreed there was significant CAD at ICA, intermediate when there was disagreement (3.1%), and lowest when they agreed there was no significant CAD (0.9%) (5).

Limitations of this study should be considered. We cannot know whether the 41% fewer interpretations of significant CAD would have translated into fewer referrals to ICA and revascularization. Further, in patients in whom the site readers found significant CAD but the core laboratory readers did not, we cannot know whether management based on this finding including medical treatment and/or revascularization might have prevented myocardial infarction or cardiovascular death. However, randomized trials have demonstrated that coronary revascularization in stable chest pain does not prevent myocardial infarction or cardiovascular death, except in those with left main, proximal left anterior descending, or three-vessel CAD (28).

In conclusion, core laboratory interpretation of coronary CT angiography classified substantially fewer patients as having significant CAD compared with site interpretation. The results suggest an opportunity for improvement in routine clinical interpretation of coronary CT angiography.

Disclosures of Conflicts of Interest: M.T.L. Activities related to the present article: received support from the American Roentgen Ray Society Scholarship. Activities not related to the present article: is a consultant to PQBypass. Other relationships: disclosed no relevant relationships. **N.M.M.** Activities related to the present article: received support from the National Heart, Lung and Blood Institute (T32 grant HL076136). Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **T.M.** disclosed no relevant relationships. **D.O.B.** disclosed no relevant relationships. **H.E.** disclosed no relevant relationships. **S.B.P.** disclosed no relevant relationships. **B.F.** disclosed no relevant relationships. **M.E.M.** disclosed no relevant relationships. **S.H.** disclosed no relevant relationships. **C.Y.** disclosed no relevant relationships. **S.A.** disclosed no relevant relationships. **Q.A.T.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received grants from Ziosoft; is a consultant to HeartFlow. Other relationships: disclosed no relevant relationships. **B.B.G.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received research grants and consulting fees from Siemens Healthcare. Other relationships: disclosed no relevant relationships. **M.R.P.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received grants from HeartFlow, Volcano, Janssen, Johnson & Johnson, Astra Zeneca, and Agency for Healthcare Research & Quality; personal fees for advisory boards from Astra Zeneca, Bayer, and Otsuka. Other relationships: disclosed no relevant relationships. **M.F.** Activities related to the present article: received support from the American Heart Association Fellow to Faculty Award. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **P.S.D.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received grants from HeartFlow and GE Healthcare. Other relationships: disclosed no relevant relationships. **U.H.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: received grants from HeartFlow and Siemens Healthcare; is a consultant to HeartFlow.

References

1. Hoffmann U, Ferencik M, Udelson JE, et al. Prognostic value of noninvasive cardiovascular testing in patients with stable chest pain: insights from the PROMISE trial (Prospective Multicenter Imaging Study for Evaluation of Chest Pain). Circulation 2017;135(24):2320–2332.
2. Douglas PS, Hoffmann U, Patel MR, et al. Outcomes of anatomical versus functional testing for coronary artery disease. N Engl J Med 2015;372(14):1291–1300.
3. Douglas PS, Hoffmann U, Lee KL, et al. Prospective Multicenter Imaging Study for Evaluation of Chest Pain: rationale and design of the PROMISE trial. Am Heart J 2014;167(6):796–803.e1.

4. Begg CB, McNeil BJ. Assessment of radiologic tests: control of bias and other design considerations. Radiology 1988;167(2):565–569.

5. Shah R, Yow E, Jones WS, et al. Comparison of visual assessment of coronary stenosis with independent quantitative coronary angiography: findings from the Prospective Multicenter Imaging Study for Evaluation of Chest Pain (PROMISE) trial. Am Heart J 2017;184:1–9.

6. Chakrabarti AK, Grau-Sepulveda MV, O'Brien S, et al. Angiographic validation of the American College of Cardiology Foundation: the Society of Thoracic Surgeons Collaboration on the Comparative Effectiveness of Revascularization Strategies study. Circ Cardiovasc Interv 2014;7(1):11–18.

7. Brener SJ, Cristea E, Lansky AJ, Fahy M, Mehran R, Stone GW. Operator versus core laboratory assessment of angiographic reperfusion markers in patients undergoing primary percutaneous coronary intervention for ST-segment-elevation myocardial infarction. Circ Cardiovasc Interv 2012;5(4):563–569.

8. Abbara S, Arbab-Zadeh A, Callister TQ, et al. SCCT guidelines for performance of coronary computed tomographic angiography: a report of the Society of Cardiovascular Computed Tomography guidelines committee. J Cardiovasc Comput Tomogr 2009;3(3):190–204.

9. Budoff MJ, Cohen MC, Garcia MJ, et al. ACCF/AHA clinical competence statement on cardiac imaging with computed tomography and magnetic resonance: a report of the American College of Cardiology Foundation/American Heart Association/American College of Physicians task force on clinical competence and training. J Am Coll Cardiol 2005; 46(2):383–402.

10. Pelberg R, Budoff M, Goraya T, et al. Training, competency, and certification in cardiac CT: a summary statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr 2011;5(5):279–285.

11. Raff GL, Abidov A, Achenbach S, et al. SCCT guidelines for the interpretation and reporting of coronary computed tomographic angiography. J Cardiovasc Comput Tomogr 2009;3(2):122–136.

12. Williams MC, Golay SK, Hunter A, et al. Observer variability in the assessment of CT coronary angiography and coronary artery calcium score: substudy of the Scottish COmputed Tomography of the HEART (SCOT-HEART) trial. Open Heart 2015;2(1):e000234.

13. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982;247(18):2543–2546.

14. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15(4):361–387.

15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.

16. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease: results from the prospective multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) trial. J Am Coll Cardiol 2008;52(21):1724–1732.

17. Meijboom WB, Meijs MF, Schuijf JD, et al. Diagnostic accuracy of 64-slice computed tomography coronary angiography: a prospective, multicenter, multivendor study. J Am Coll Cardiol 2008;52(25):2135–2144.

18. Herzog C, Kerl JM, De Rosa S, et al. Influence of observer experience and training on proficiency in coronary CT angiography interpretation. Eur J Radiol 2013;82(8):1240–1247.

19. Ovrehus KA, Munkholm H, Bøttcher M, Bøtker HE, Nørgaard BL. Coronary computed tomographic angiography in patients suspected of coronary artery disease: impact of observer experience on diagnostic performance and interobserver reproducibility. J Cardiovasc Comput Tomogr 2010;4(3):186–194.

20. Pugliese F, Hunink MG, Gruszczynska K, et al. Learning curve for coronary CT angiography: what constitutes sufficient training? Radiology 2009;251(2):359–368.

21. Ryan T, Berlacher K, Lindner JR, Mankad SV, Rose GA, Wang A. COCATS 4 task force 5: training in echocardiography. J Am Coll Cardiol 2015;65(17):1786–1799.

22. Dilsizian V, Arrighi JA, Cohen RS, Miller TD, Solomon AJ, Udelson JE. COCATS 4 task force 6: training in nuclear cardiology. J Am Coll Cardiol 2015;65(17):1800–1809.

23. Taylor AJ, Patrick J, Abbara S, et al. Relationship between previous training and experience and results of the certification examination in cardiovascular computed tomography. JACC Cardiovasc Imaging 2010;3(9):976–980.

24. Rana JS, Tabada GH, Solomon MD, et al. Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary, multiethnic population. J Am Coll Cardiol 2016; 67(18):2118–2130.

25. Ridker PM, Rifai N, Rose L, Buring JE, Cook NR. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. N Engl J Med 2002;347(20):1557–1565.

26. den Dekker MA, de Smet K, de Bock GH, Tio RA, Oudkerk M, Vliegenthart R. Diagnostic performance of coronary CT angiography for stenosis detection according to calcium score: systematic review and meta-analysis. Eur Radiol 2012;22(12):2688–2698.

27. Budoff MJ, Mayrhofer T, Ferencik M, et al. The prognostic value of coronary artery calcium in the PROMISE Study. Circulation 2017 Aug 28. [Epub ahead of print].

28. Boden WE, O'Rourke RA, Teo KK, et al. Optimal medical therapy with or without PCI for stable coronary disease. N Engl J Med 2007;356(15):1503–1516.