European Society of Cardiology Recommended CAD Consortium Pre-Test Probability Scores More Accurately Predict Obstructive Coronary Disease and Cardiovascular Events Than the Diamond and Forrester Score: The Partners Registry

Running title: Bittencourt et al.; Pretest probability of obstructive coronary artery disease

Marcio Sommer Bittencourt, MD, MPH, PhD1-3*; Edward Hulten, MD, MPH2,4*; Tamar S. Polonsky, MD, MSCI5; Udo Hoffman, MD, MPH6; Khurram Nasir, MD, MPH7; Suhny Abbara, MD6,8; Marcelo Di Carli, MD2; Ron Blankstein, MD2

1Center for Clinical and Epidemiological Research, University Hospital & São Paulo state cancer institute; University of São Paulo School of Medicine, Sao Paulo, Brazil; 2Cardiovascular Imaging Program, Departments of Medicine and Radiology; Brigham and Women’s Hospital; Harvard Medical School, Boston, MA; 3Preventive Medicine Center, Hospital Israelita Albert Einstein, São Paulo, Brazil; 4Cardiology Service, Department of Internal Medicine, Walter Reed National Military Medical Center, Bethesda, MD; 5Department of Medicine, University of Chicago; 6Cardiac MR PET CT Program, Department of Radiology, Division of Cardiac Imaging, Massachusetts General Hospital; Harvard Medical School, Boston, MA; 7Baptist Health South Florida, Miami, FL; 8Department of Radiology, University of Texas Southwestern, Dallas, TX

* Drs. Bittencourt and Hulten contributed equally as first author

Address for Correspondence:
Ron Blankstein, MD
Cardiovascular Division and Department of Radiology
Brigham and Women’s Hospital
75 Francis St, Boston, MA 02115
Tel: (857) 307-1989
Fax: (857) 307-1955
Email: rblankstein@partners.org

Journal Subject Terms: Prognosis; Diagnostic Testing
Abstract

**Background**—The most appropriate score for evaluating the pretest probability of obstructive coronary artery disease (CAD) is unknown. We sought to compare the Diamond Forrester (DF) score with the two CAD consortium scores recently recommended by the European Society of Cardiology.

**Methods**—We included 2274 consecutive patients (age 56±13 years, 57% males) without prior CAD referred for coronary CT angiography (CTA). CTA findings were used to determine the presence or absence of obstructive CAD (≥50% stenosis). We compared the DF score with the two CAD consortium scores with respect to their ability to predict obstructive CAD as well as the potential implications of these score on downstream utilization of testing for CAD, as recommended by current guidelines.

**Results**—The DF score did not satisfactorily fit the data, and resulted in a significant overestimation of the prevalence of obstructive CAD (p<0.001), the CAD consortium basic had no significant lack of fitness, while the CAD consortium clinical provided adequate goodness-of-fit (p=0.39). The DF had a lower discrimination for obstructive CAD, with an area under the ROC curve of 0.713 vs. 0.752 and 0.791 for the CAD consortium models (p<0.001 for both). Consequently, the use of the DF score was associated with fewer individuals being categorized as requiring no additional testing (8.3%) when compared to the CAD consortium models (24.6% and 30.0%, p<0.001). The proportion of individuals with a high pretest probability was 18% with the DF and only 1.1% with the CAD consortium scores (p<0.001)

**Conclusions**—Among contemporary patients referred for non-invasive testing, the DF risk score over estimates the risk of obstructive CAD. On the other hand, the CAD consortium scores offered improved goodness-of-fit and discrimination, and thus their use could decrease the need for non-invasive or invasive testing, while increasing the yield of such tests.

**Key words:** stable coronary artery disease; chest pain; risk prediction; prognosis
Clinical Perspective

What is new?
- Among symptomatic individuals referred for coronary computed tomography angiography, the CAD consortium clinical pretest probability score demonstrated improved calibration and discrimination for the prediction of obstructive coronary artery disease than the Diamond and Forrester classification.
- When applying these observed differences in pre-test probability of obstructive CAD to guidelines-based patient management algorithms, we projected that the use of the newer score could decrease the proportion of individuals in whom testing (both non-invasive or invasive) would be recommended, and increase the yield of diagnosing obstructive CAD.

What are the clinical implications?
- All scoring systems used in our study were based on clinically available information, and changing from one score to another can be easily accomplished.
- Our results support using the CAD consortium clinical score instead of the Diamond and Forrester classification. This change will likely result in a reduction in the overall need for additional testing as a larger number of individuals will be appropriately reclassified at low probability of obstructive CAD.
Individuals with stable coronary artery disease (CAD) experience lower quality of life and higher rates of adverse cardiovascular (CV) events and mortality compared to healthy individuals.\(^1\) In individuals with chest pain, both non-invasive and invasive investigations of CAD are often used to establish prognosis as well as guide treatment.\(^2,\)\(^3\) However, studies have shown a relatively low prevalence of either ischemia or obstructive CAD on noninvasive imaging and invasive angiography in this population.\(^4,\)\(^5\) Therefore, additional methods are needed to improve patient selection for such testing.

U.S. and European guidelines recommend using a diagnostic strategy tailored to the individual’s pretest probability of obstructive CAD.\(^6,\)\(^7\) While individuals with a very low probability may not need further investigation, among those with an extremely high pretest probability it may be reasonable to proceed directly to invasive angiography (IA) for risk stratification.\(^6,\)\(^7\) For individuals with an intermediate probability of obstructive CAD, guidelines recommend further evaluation with non-invasive CV imaging.\(^6,\)\(^7\)

The first score to calculate the pretest probability of obstructive CAD, introduced over three decades ago in a seminal work by Diamond and Forrester (DF),\(^8\) is a simple and easy score which is recommended in the American College of Cardiology/American Heart Association (ACC/AHA) guidelines and appropriate use criteria for stable CAD.\(^6,\)\(^9-12\) However, a recent European study demonstrated that although the predictors selected by Diamond and Forrester are robust, their calibration was not adequate for the modern population of patients investigated for CAD.\(^13,\)\(^14\) Recognizing these limitations, the most recent European Society of Cardiology (ESC) guidelines for stable CAD have replaced the DF score with two new revised scores.\(^7\)

As a result, significant differences now exist between European and U.S. guidelines for evaluating individuals with chest pain. Yet, the implications resulting from these differences
have not been evaluated. Moreover, no study to date has evaluated the potential role of these clinical scores to predict adverse CV events. In the present study, we compared the ACC/AHA recommended DF score with the two scores recommended by the ESC in respect to their ability to predict the presence of obstructive CAD, as well as the association of these scores with adverse CV events in a cohort of individuals with no prior CAD referred for coronary computed tomography angiography (CTA).

Methods

Study population

We included in a registry all consecutive subjects 18 years or older who underwent a clinically indicated coronary CTA for the evaluation of suspected CAD at the Massachusetts General Hospital or Brigham and Women’s Hospital from 2004 to 2011.\textsuperscript{15} We excluded patients who were missing any of the clinical information needed to calculate the pre-test probability, who had non-diagnostic coronary CTA images or who had incomplete follow up information. Patients with congenital heart disease, heart transplant or prior CAD, defined as prior percutaneous coronary interventions (PCI), coronary artery bypass surgery (CABG) or MI, were also excluded (supplemental figure 1). The Human Research Committee of both institutions approved the study. Retrospective review of previously existing clinical data was considered exempt by the Partners IRB in accordance with standard practice. IRB permission was granted to contact patients, if indicated, by mail or phone who then voluntarily provided consent at the time of contact.

Ascertainment of Risk Factors

Systemic arterial hypertension was defined as a systolic blood pressure $> 140$ mmHg, diastolic
blood pressure > 90 mmHg, or diagnosis/treatment of hypertension. Dyslipidemia was defined as total cholesterol > 240 mg/dL or serum triglycerides > 150 mg/dL or high-density lipoprotein cholesterol (HDL) <40 mg/dL (men) or < 50 mg/dL (women) or diagnosis/treatment of dyslipidemia. Diabetes was defined by a hemoglobin A1C ≥6.5%16, physician-based diagnosis, or use of glucose-lowering medications. Smoking was defined as current (tobacco products used within the last month), former or never. Family history of premature CAD was defined as a self-reported history of any first-degree family member with a history of MI PCI or revascularization prior to age 60. Ethnicity was self-reported and identified as Caucasian, African American, Hispanic, Asian, Other, unknown or refused.

CTA exam acquisition and interpretation

All scans were performed using 64 row computed tomography scanners or newer technologies. The studies were performed according to established guidelines,17, 18 and institutional protocols at the time of the scan. Following each scan, the images were reconstructed in single or multiphase datasets and images were interpreted using axial and multiplanar reformations.

All scans were analyzed by level III trained cardiologists or radiologists with extensive experience in coronary CTA analysis. The coronary CTAs were interpreted according to current guidelines18 using a previously published 18-segment model.18 Each coronary segment with a greater than 2 mm diameter was analyzed for the presence of coronary atherosclerosis and each lesion was quantified by visual estimation into three categories: normal, non-obstructive disease (1 to 49% stenosis) and obstructive disease (≥50% stenosis). Obstructive CAD was defined as at least one segment with a lesion with ≥50% stenosis.

Pre-test probability scores

The DF pretest probability was calculated based on the chest pain type (nonanginal chest pain,
atypical angina or typical angina), gender and age. This score was developed using conditional probability and modeled to predict a lesion ≥50% stenosis for individuals between the ages of 30 to 69 years. We used a revised version of the DF score, which allows inclusion of patients older than 69 and incorporates age as a continuous variable (table 1).

We compared the DF score to two CAD consortium models according to the coefficients provided by Genders, et al. The first model, named CAD consortium basic, was based on age, sex and chest pain type (typical, atypical or nonanginal chest pain). Although the score uses the same parameters as the original DF, the model was developed to detect a ≥50% stenosis on invasive angiography or coronary CTA. The score was developed using more advance statistical modeling strategies, which were not available when the DF model was derived. Additionally, the population had a lower prevalence of disease than the original DF derivation cohort did (table 1).

The second model, named CAD consortium clinical, included the same characteristics as the CAD basic, but also included the following clinical risk factors: diabetes, smoking status, hypertension and dyslipidemia (table 1). Moreover, for the clinical model, the presence of typical chest pain was weighted less in diabetics vs. non-diabetics. The CAD consortium clinical was developed using logistic regression, and was validated for populations with a lower prevalence of disease than the population used in the derivation of DF. Prior studies have provided detailed definitions and validation for each parameter in the model.

**Cardiovascular outcomes**

For the survival analysis we used the primary composite end-point of the major adverse cardiovascular events (MACE) composed of CV mortality, non-fatal MI, late coronary revascularization (>90 days) and unstable angina requiring hospital admission.

All patient charts were reviewed for the adjudication of CV events by two cardiologists
who were blinded to coronary CTA results. To ensure that events outside of our healthcare network were captured, a standardized questionnaire was mailed to each patient. Additionally, patients had the option of completing a web-based version of the questionnaire via the REDCap (Research Electronic Data Capture) system,19 which is encrypted, secure, and HIPAA compliant. For patients who did not reply to the questionnaire upon repeated mailings, scripted phone interviews were performed based on the questionnaire. All self-reported events were verified via outside medical record review by two cardiologists blinded to coronary CTA results with discordant events adjudicated by consensus.

Deaths were confirmed by the Social Security Death Index. For all patients who died the cause of death was obtained from the National Death Index. When data were not available, records from the Massachusetts Department of Vital Statistics were obtained. In addition, other pertinent clinical records (e.g. death notes, autopsy findings, hospice notes) related to the cause of death were reviewed. Using all available data, the cause of death for each patient was adjudicated by two cardiologists blinded to the coronary CTA results. The cause of death was considered to be of cardiovascular origin if the primary cause was defined due to acute MI, atherosclerotic coronary vascular disease, congestive heart failure, valvular heart disease, arrhythmias, stroke, or other structural or primary cardiac cause of death. MI was defined when at least two of the following three criteria were met: chest pain or equivalent symptom complex; positive cardiac biomarkers; or typical electrocardiogram (ECG) changes.20 For revascularizations, the time to the first coronary revascularization procedure (PCI or CABG) was evaluated. Early revascularizations (≤90 days post coronary CTA) were censored in the survival analysis to minimize verification bias,21-23 as patients with ≥50% stenosis by coronary CTA may be referred for invasive angiography and revascularization based on the coronary CTA results.
alone. On the other hand, late revascularizations (>90 days post coronary CTA) are more likely to be associated with CAD progression, and were therefore included as part of the composite end-point. Unstable angina requiring admission was defined as chest pain or chest pain equivalent with dynamic ECG changes such as ST depression or T wave inversion, but without abnormal cardiac biomarkers, and characterized by: 1) rest symptoms; 2) new onset angina (less than 2 months duration); or, 3) increasing duration or severity of previously stable anginal symptoms.24

**Planned strategy according to guidelines**

To evaluate how the results of each score might influence use of downstream non-invasive and invasive testing, we stratified each score result as low (<5%), intermediate (5 – 70%) or high (>70%) pretest probability of obstructive CAD, defined as a ≥50% stenosis in at least one vessel. Those cut-off values were based on the 2012 ACC/AHA guidelines which state that individuals with a low (<5%) pretest probability are unlikely to benefit from additional testing, those with an intermediate probability (5 – 70%) are most likely to benefit from an initial non-invasive test. Lastly, individuals with a high pretest probability might be considered to have presumed CAD, though additional testing, including direct referral for IA, should be considered for further risk stratification.

**Statistical analysis**

Continuous variables are expressed as mean ± standard deviation, except for time of follow up, which is expressed as median and interquartile range. Categorical variables are presented as absolute numbers and frequencies. Differences between groups were tested using chi-square or Fisher’s exact tests for discrete variables and one-way analysis of variance for continuous variables.
For each of the pretest probability scores (DF, CAD consortium basic and CAD consortium clinical) we plotted the pretest probability across deciles of the population vs. the actual observed presence of obstructive disease on coronary CTA. Additionally, we calculated the Hosmer-Lemeshow goodness-of-fit for each of the prediction models. In this analysis, models with a p<0.05 were considered to have inadequate fit of the data, while higher p values signified adequate fit. To evaluate model discrimination we constructed ROC curves, and compared the area under the ROC curve (AUC) for each model, using the presence of obstructive (>50% stenosis) as the outcome. The positive predictive value of each score was defined as the proportion of patients with obstructive CAD among those classified as high pretest probability by the score, whereas the false negative rate was defined as the proportion of individuals with obstructive CAD among those classified as low pretest probability by each score.

In order to evaluate the ability to predict future MACE, we built univariable Cox proportional hazard models for each of the three scores, constructed ROC curves and compared the AUC for the three models using the Harrel’s c-index, and calculating the confidence intervals using Somers D statistics. Additionally, we evaluated the goodness-of-fit using the Gronnesby and Borgan test, where a non-significant p value indicates adequate fit of the model. The assumption of proportional hazards was tested using a formal significance test based on the unscaled and scaled Schoenfeld residuals and resulted in non-significant findings in in all analyses.

To compare the prevalence of individuals stratified as low (<5%), intermediate (5 – 70%) or high (>70%) pretest probability of obstructive CAD, define as a ≥50% stenosis in at least one vessel, we performed chi-square tests. Statistical analysis was performed using Stata version 12
(Statacorp, College Station, USA), and statistical significance was defined as a two-tailed p
<0.05.

**Results**

**Patient Population and baseline characteristics**

Among 2274 patients who met our inclusion and exclusion criteria and had all clinical
information needed to calculate all scores, the mean age was 56±13 years, 1289 (57%) were
males, and 501 (22%) had obstructive CAD. Other baseline characteristics are presented in table
2. Information on the differences between included and excluded individuals is presented in
supplemental table 1.

When stratified by the presence or absence of obstructive CAD on coronary CTA, the
presence of obstructive disease was associated with older age, male sex, symptoms and all
traditional CV risk factors, except for family history of premature CAD (table 2). The pre-test
probability of CAD of individuals with obstructive disease was higher than for those with no
obstructive disease for all three scores (DF, CAD consortium basic and CAD consortium
clinical) (table 2).

**Comparison of pre-test probability scores to predict obstructive CAD**

While higher scores were associated with a higher probability of obstructive disease for all three
scores (figure 1), the DF score demonstrated a poor model fit (p<0.001), leading to an important
overestimation of disease prevalence, mainly for individuals with higher scores. For example, in
the second blue dot from right to left in figure 1A the pre-test score estimated a probability of
disease of 0.70, although the actual observed prevalence of obstructive CAD was only 0.30
based on coronary CTA. For the CAD consortium basic no significant lack of fitness was noted,
though small deviations cannot be excluded due to borderline p value of 0.08, while the CAD consortium clinical provided adequate goodness-of-fit (p=0.39). (figures 1B and 1C).

The DF score was also less capable of discriminating between individuals with or without obstructive CAD. The DF had a lower discrimination than the other models, while CAD consortium clinical had the highest discrimination of all three. (Figure 2).

Comparison of pre-test probability scores to predict adverse CV events

During a median follow up of 3.3 years (interquartile range: 1.9 – 4.7) the primary outcome of MACE occurred in 148 (6.5%) individuals. This included 38 (1.7%) cardiovascular deaths, 33 (1.5%) myocardial infarctions, 27 (1.3%) unstable angina cases, and 67 (3%) late revascularizations.

Although none of these scores was designed to predict future CV events, we have compared their ability to discriminate between individuals who experienced incident MACE from those who did not. The DF score had the worst discriminatory ability for MACE (AUC: 0.623 [95%CI: 0.578-0.668]), the CAD consortium basic had a significantly higher AUC (0.638 [95%CI: 0.593 – 0.682), while the CAD consortium clinical had the highest AUC (0.687 [95%CI: 0.646 – 0.728).

Implications of pre-test probability scores on utilization of cardiac testing

With the use of DF, only 188 individuals (8.3% of the entire cohort) were classified as low risk and thus would not need additional testing according to guidelines. This proportion increased to 560 (24.6%, p<0.001) with the use of the CAD consortium basic model; and to 6823 (30.0%, p<0.001) with the use of the CAD consortium clinical model (figure 3). The use of the DF score also resulted in a significantly larger proportion of individuals with a high pre-test probability compared to the CAD consortium scores: (410 individuals, 18.0% vs 1.1%. For each of the CAD
consortium scores (figure 3).

Importantly, while the positive predictive ratio was only 44% (180/410) when using DF, it was higher for both the CAD consortium basic (71%, 17/24) and the CAD consortium clinical (79%, 19/24) p<0.001 for both when compared to DF.

Notably, with all three scores the false negative rate was low. There were only 9 (0.3%) individuals with obstructive disease who were inappropriately classified as having a low pre-test probability by DF, 30 (1.3%, p<0.001) by the CAD consortium basic, and to 21 (0.9%, p<0.001) by the CAD consortium clinical score.

When examining the rate of MACE among individuals categorized as low risk by the various risk score, we observed a low event rate for all scores (0.74% per year for DF, 0.96% per year for the CAD consortium basic and 0.59% per year for the CAD consortium clinical (p not significant for all pairwise comparisons).

Discussion
In the present study, we have demonstrated that estimation of the pretest probability of CAD as recommended by the ACC/AHA guidelines or by the ESC guidelines results in significantly different risk estimations, which may influence the utilization of downstream testing and medical therapies for a significant proportion of individuals evaluated for chest pain. The Diamond and Forrester score resulted in a significant overestimation of the prevalence of obstructive CAD when compared to the two models recommended by the ESC. As a result, the use of this risk score may result in overutilization of non-invasive and invasive diagnostic testing in individuals with a low prevalence of disease. In addition, the DF had a lower discrimination for obstructive disease when compared to the ESC recommended scores.
Our results suggest that due to the differences in risk classification observed in our study, replacing the DF risk score by the CAD consortium clinical score would result in as much as a 95% reduction in the number of individuals categorized as having a high pretest probability. Since this is a sub-group who, per guideline recommendations, would subsequently be treated as having presumed CAD and could be referred directly to invasive angiography, a reduction in this sub-group could significantly decrease utilization of testing in a group which is unlikely to have significant disease. At the same time, our study estimates that using the CAD consortium clinical score could result in a three-fold increase in the number of individuals who would be categorized as low risk and subsequently would not require any additional testing (from 9% to 31%). Moreover, the use of the CAD consortium clinical score would increase the yield of testing (i.e. proportion of individuals referred for testing who are found to have abnormal results) among those with a high pretest probability from 44% to about 80%. Interestingly, although none of these scores were designed to predict incident CV events, the CAD consortium scores recommended by the ESC also had a significantly better discrimination for incident CV events.

The DF risk score represents a seminal achievement in clinical reasoning and has built the foundation for assessment of pretest probability of CAD using Bayesian reasoning. Nonetheless, the original validation of the DF risk score occurred decades ago among patients referred for IA or upon autopsy, while contemporary patients are more likely to undergo non-invasive testing and may be of lower probability of disease. There are several factors that may account for this trend: (a) wider availability and utilization of non-invasive testing; (b) increased awareness and recognition by patients of anginal symptoms and thus higher likelihood of presenting earlier in the disease course; (c) increased use of preventive medical therapies even before the clinical presentation starts. Our results are consistent with several studies investigating
the accuracy of the DF risk score in the current era. One important study occurred in the international Coronary CT Angiography Evaluation For Clinical Outcomes: An International Multicenter (CONFRIM) Registry. This study evaluated 14,048 subjects with suspected CAD who underwent CTA and determined that DF markedly overestimated the likelihood of obstructive CAD at all centers in the cohort and across all age and gender groups. For example, patients presenting with atypical angina had an observed prevalence of 15% obstructive CAD versus 47% predicted. Similarly, those with typical angina had 29% observed versus 86% predicted probability. This overestimation was also demonstrated in other studies.

Emphasizing the implications of risk scores on the appropriate use criteria, a study by Wasfy, et al. suggested that the choice of pretest probability scores could reclassify the level of appropriateness in a significant proportion of patients referred for non-invasive testing for CAD symptoms. This study concluded that the Duke clinical score was more accurate to predict obstructive CAD, and that this could significantly affect the classification of test appropriateness based on published criteria, however their cohort was much smaller than the current study and the data precedes the development of the CAD consortium scores.

Based on these limitations of the DF scores, the CAD consortium developed and validated new clinical risk scores which were based not only upon patients referred to IA, who are likely to have higher pretest probabilities, but also based upon low and intermediate pretest probability individuals who are referred to coronary CTA. Both scores developed and validated by this consortium were based on multicenter registries of patients in the U.S. and Europe. Although their study performed internal validation of the scores, to our knowledge, the present data is the first external validation of their scores in the literature, and the results are remarkably similar. Despite differences in data collection, secular trends, local patterns for referrals and
other potential differences between studies, our findings support the CAD consortium scores as extremely robust scores for the prediction of obstructive disease. These scores were better calibrated than DF in our sample, with notable improvement in goodness-of-fit for the CAD consortium clinical score. Additionally, they had a significantly better discrimination for obstructive disease when compared to the DF score.

Our finding that contemporary risk scores classify fewer patients as high and more patients as low pretest probability carries substantial economic implications. While less than 10% of our cohort would be considered of low enough pretest probability to defer additional testing by DF, more than 30% could potentially defer further testing by the CAD consortium scores. Likewise, a sizeable reduction in the individuals who would be presumed to have obstructive CAD and might be considered candidates for invasive risk assessment, from 16.3% to 1% of the population, was noted. Reducing over-estimation of pre-test risk among individuals being evaluated for potential testing could lead to a reduction in non-invasive and invasive testing. In the current era of cost containment, and given the low prevalence of obstructive CAD among patients referred for invasive and non-invasive CAD testing, even if the reduction in further testing is smaller than anticipated by the above estimates, it could still have important economic implications.

Another important finding of our study was the evaluation of prognosis according to the pretest probability scores. Our study suggests that although the CAD consortium clinical had a somewhat better discrimination for future CV events when compared to both the DF and the CAD consortium basic, none of the models had a good discrimination for future cardiovascular events. Whilst this supports this score as the most appropriate pretest probability score for the investigation of stable CAD, this was not unexpected, as the main difference between this score
and the other two is that it also incorporates clinical risk factors for CAD (such as diabetes, hypertension, smoking and dyslipidemia), which are well known markers of future CV events. Nevertheless, a key finding on prognosis in the present study is the fact that irrespective of the pretest probability score used, the rate of MACE in individuals with a low pretest probability score is equally low. This finding further supports the safety of withholding additional testing in this population.

Our study must, however, be read within the context of its design. First, the AHA/ACC guidelines suggest that either the DF or the Duke Clinical Scores can be used to estimate the pretest probability score. While we did not have all the clinical information, such as ECG findings, to calculate the Duke Clinical Score, prior data suggests both studies have comparable performance. Additionally, although the present findings corroborate data from other studies and validates previously developed scores, the actual reductions in the use of additional testing may be highly variable and depends on local clinical practice patterns. While our analysis assumes that patient management will be dictated by guidelines, the real life impact of these scores depends on the actual implementation of those recommendations in actual practice as well as variability in physician and patient preferences. Furthermore, we used coronary CTA as the reference standard to establish the diagnosis of obstructive CAD. Since coronary CTA may overestimate the prevalence of obstructive disease, use of invasive angiography could have resulted in an even lower observed prevalence of disease (and possibly need for additional testing) than predicted by our study. Additionally, Hosmer-Lemeshow goodness-of-fit has known limitations for the evaluation of calibration. Also, while our results are most applicable to patients referred for coronary CTA, we anticipate that these findings would have important implications for patients who are considered for other non-invasive or invasive testing, since (a) the absence of
obstructive CAD on coronary CTA has a very high negative predictive value to exclude ischemia on functional testing\textsuperscript{30}, and (b) the consortium clinical score was validated both for individuals undergoing coronary CTA and for those undergoing IA. Nevertheless, our results may be less applicable to higher risk cohorts and should not be extrapolated to the emergency department setting.

In conclusion, use of the CAD consortium pretest probability scores, particularly the CAD consortium clinical score, results in improved model goodness-of-fit and provides better discrimination for the detection of obstructive CAD than the DF risk score. Although warranting further validation using both invasive and non-invasive testing, our results, which are based on coronary CTA, suggest that use of the CAD consortium scores could potentially reduce unnecessary referrals for non-invasive CV imaging tests and invasive angiography for the investigation of stable chest pain. While requiring additional external validation, our findings suggest that such a reduction in the testing may also be accompanied by an increase in the yield of both non-invasive and invasive testing. Thus, the current study supports the replacement of the DF risk score by the CAD consortium clinical score as the recommended pretest probability score for the investigation of patients with suspected obstructive CAD.

**Sources of Funding:** Dr Bittencourt was supported by the J.P. Lemann Foundation as a Jorge Paulo Lemann Harvard Medical School Cardiovascular Fellow at Brigham and Women’s Hospital.

**Conflict of Interest Disclosures:** None.

**References**


9. Taylor AJ, Cerqueira M, Hodgson JM, Mark D, Min J, O’Gara P, Rubin GD, Kramer CM, Min JK, Patel MR, Peterson ED, Wolk MJ, Allen JM, Bailey S, Douglas PS, Hendel RC, Patel MR, Shaw L, Stainback RF. Accf/scct/acep/aha/ascn/nasci/scai/scct/scmr/snm 2009 appropriate use criteria for cardiac radionuclide imaging: A report of the american college of cardiology foundation appropriate use criteria task force, the society for cardiovascular magnetic resonance, the society for cardiovascular computed tomography, the society for nuclear medicine and molecular imaging, the american college of radiology, the american heart association, the american society of echocardiography, the american society of nuclear cardiology, the society of nuclear medicine, the society of nuclear medicine and molecular imaging, and the society for cardiovascular magnetic resonance. *Circulation.* 2009;119:e561-587


21


score to predict obstructive coronary artery disease by computed tomographic angiography. *Am J Cardiol*. 2012;109:998-1004


Table 1: Information on how Diamond Forrester, CAD consortium basic and clinical were derived

<table>
<thead>
<tr>
<th>Risk Scores</th>
<th>Diamond Forrester</th>
<th>CAD consortium basic</th>
<th>CAD consortium clinical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended by</td>
<td>ACC / AHA</td>
<td>ESC</td>
<td>ESC</td>
</tr>
<tr>
<td>Includes</td>
<td>Age, gender, angina typicality</td>
<td>Age, gender, angina typicality</td>
<td>Age, gender, angina typicality, diabetes, smoking status, hypertension and dyslipidemia</td>
</tr>
<tr>
<td>Validated by</td>
<td>Patients referred to invasive angiography during the 1970s</td>
<td>Contemporary patients referred to invasive angiography or coronary CTA</td>
<td>Contemporary patients referred to invasive angiography or coronary CTA</td>
</tr>
</tbody>
</table>
Table 2: Baseline demographic characteristics according to the presence and severity of CAD.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>&lt;50% Stenosis</th>
<th>≥50% Stenosis</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>2274</td>
<td>1773 (78%)</td>
<td>501 (22%)</td>
<td></td>
</tr>
<tr>
<td>Age*</td>
<td>56±13</td>
<td>53±13</td>
<td>64±10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Male (%)</td>
<td>1289 (57%)</td>
<td>919 (52%)</td>
<td>370 (74%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Caucasian</td>
<td>1844 (81%)</td>
<td>1409 (79%)</td>
<td>435 (87%)</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>128 (6%)</td>
<td>118 (7%)</td>
<td>10 (6%)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>101 (4%)</td>
<td>89 (5%)</td>
<td>12 (2%)</td>
<td></td>
</tr>
<tr>
<td>Other ethnicity</td>
<td>87 (4%)</td>
<td>69 (4%)</td>
<td>18 (4%)</td>
<td></td>
</tr>
<tr>
<td>Unknown or refused</td>
<td>114 (5%)</td>
<td>88 (5%)</td>
<td>26 (5%)</td>
<td></td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>1207 (53%)</td>
<td>822 (46%)</td>
<td>385 (77%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes Mellitus (%)</td>
<td>354 (16%)</td>
<td>212 (13%)</td>
<td>131 (26%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dyslipidemia (%)</td>
<td>1251 (55%)</td>
<td>875 (48%)</td>
<td>406 (81%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Prior smoking</td>
<td>653 (29%)</td>
<td>458 (26%)</td>
<td>195 (39%)</td>
<td></td>
</tr>
<tr>
<td>Current smoking</td>
<td>564 (10%)</td>
<td>172 (10%)</td>
<td>59 (12%)</td>
<td></td>
</tr>
<tr>
<td>Family history of CAD</td>
<td>654 (29%)</td>
<td>501 (28%)</td>
<td>153 (31%)</td>
<td>0.55</td>
</tr>
<tr>
<td>Symptoms (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Non-anginal CP</td>
<td>1066 (47%)</td>
<td>826 (47%)</td>
<td>240 (48%)</td>
<td></td>
</tr>
<tr>
<td>Atypical CP</td>
<td>1003 (44%)</td>
<td>797 (45%)</td>
<td>206 (41%)</td>
<td></td>
</tr>
<tr>
<td>Typical CP</td>
<td>205 (9%)</td>
<td>150 (8%)</td>
<td>55 (11%)</td>
<td></td>
</tr>
<tr>
<td>Diamond Forrester pre-test probability¥ (%)</td>
<td>31 (14-57)</td>
<td>26 (11-51)</td>
<td>53 (30-78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAD consortium basic pre-test probability¥ (%)</td>
<td>10 (5-18)</td>
<td>8 (4-15)</td>
<td>19 (11-32)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAD consortium clinical pre-test probability¥ (%)</td>
<td>9 (4-18)</td>
<td>7 (3-14)</td>
<td>20 (11-36)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*: results are presented as means and standard deviation; ¥: results are presented as median and quartiles.
Figure Legends:

**Figure 1**: predicted vs. observed probability of CAD for the DF (a), CAD consortium basic (b) and CAD consortium clinical (c) risk scores. The blue dots represent the deciles of pretest probability according to each scores, and the red line represents the fitted line of pretest probability to observed prevalence of obstructive CAD. The dotted line is where the line would be if agreement was perfect.

**Figure 2**: ROC curves for the prediction of obstructive CAD on coronary CTA. (p<0.001 for all comparisons)

**Figure 3**: Distribution of the population in low (green), intermediate (yellow) or high (red) pretest probability scores for the Diamond-Forrester (left), CAD consortium basic (middle) and CAD consortium clinical (right). The two CAD consortium classify significantly more individuals in the low pretest probability group and significantly fewer individuals in the high pretest probability groups. (p<0.001)
The chart shows the distribution of pretest probability categories according to different diagnostic criteria:

- **Diamond-forester**:
  - High pretest probability: 8.3%
  - Intermediate probability: 73.7%
  - Low probability: 18.0%

- **CAD consortium 2 basic**:
  - High pretest probability: 24.6%
  - Intermediate probability: 74.3%
  - Low probability: 1.1%

- **CAD consortium 2 complete**:
  - High pretest probability: 30.0%
  - Intermediate probability: 70.0%
  - Low probability: 1.1%

The categories are defined as follows:

- **High pretest probability**: Consider invasive or non-invasive testing.
- **Intermediate probability**: Non-invasive testing recommended.
- **Low probability**: No further testing required.