Nontraditional Risk Factors in Cardiovascular Disease Risk Assessment
Updated Evidence Report and Systematic Review for the US Preventive Services Task Force

Jennifer S. Lin, MD, MCR; Corinne V. Evans, MPP; Eric Johnson, PhD; Nadia Redmond, MSPH; Erin L. Coppola, MPH; Ning Smith, PhD

**IMPORTANCE** Incorporating nontraditional risk factors may improve the performance of traditional multivariable risk assessment for cardiovascular disease (CVD).

**OBJECTIVE** To systematically review evidence for the US Preventive Services Task Force on the benefits and harms of 3 nontraditional risk factors in cardiovascular risk assessment: the ankle-brachial index (ABI), high-sensitivity C-reactive protein (hsCRP) level, and coronary artery calcium (CAC) score.

**DATA SOURCES** MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials for studies published through May 22, 2017. Surveillance continued through February 7, 2018.

**STUDY SELECTION** Studies of asymptomatic adults with no known cardiovascular disease.

**DATA EXTRACTION AND SYNTHESIS** Independent critical appraisal and data abstraction by 2 reviewers.

**MAIN OUTCOMES AND MEASURES** Cardiovascular events, mortality, risk assessment performance measures (calibration, discrimination, or risk reclassification), and serious adverse events.

**RESULTS** Forty-three studies (N = 267,244) were included. No adequately powered trials have evaluated the clinical effect of risk assessment with nontraditional risk factors on patient health outcomes. The addition of the ABI (10 studies), hsCRP level (25 studies), or CAC score (19 studies) can improve both discrimination and reclassification; the magnitude and consistency of improvement varies by nontraditional risk factor. For the ABI, improvements in performance were the greatest for women, in whom traditional risk assessment has poor discrimination (C statistic change of 0.112 and net reclassification index [NRI] of 0.096). Results were inconsistent for hsCRP level, with the largest analysis (n = 166,596) showing a minimal effect on risk prediction (C statistic change of 0.0039, NRI of 0.0152). The largest improvements in discrimination (C statistic change ranging from 0.018 to 0.144) and reclassification (NRI ranging from 0.084 to 0.35) were seen for CAC score, although CAC score may inappropriately reclassify individuals not having cardiovascular events into higher-risk categories, as determined by negative nonevent NRI. Evidence for the harms of nontraditional risk factor assessment was limited to computed tomography imaging for CAC scoring (8 studies) and showed that radiation exposure is low but may result in additional testing.

**CONCLUSIONS AND RELEVANCE** There are insufficient adequately powered clinical trials evaluating the incremental effect of the ABI, hsCRP level, or CAC score in risk assessment and initiation of preventive therapy. Furthermore, the clinical meaning of improvements in measures of calibration, discrimination, and reclassification risk prediction studies is uncertain.

Published online July 10, 2018.

© 2018 American Medical Association. All rights reserved.
Cardiovascular disease (CVD) is the leading cause of death in the United States, accounting for about 1 in 3 deaths. The incidence of CVD is strongly associated with a set of traditional risk factors, which have been combined using multivariable risk assessment tools to estimate an individual’s risk for having a CVD event. However, these tools can underestimate or overestimate CVD risk. Inclusion of nontraditional biological and physiologic risk factors might improve the performance of tools. Given that risk estimates are used to guide preventive therapy such as aspirin and statins, improved risk assessment performance could result in improved CVD outcomes.

Previous research has established that abnormal values for the nontraditional risk factors of the ankle-brachial index (ABI), high-sensitivity C-reactive protein (hsCRP) level, and coronary artery calcium (CAC) score are significantly associated with morbidity and mortality. Moreover, this research has shown that these factors can improve the ability of models to distinguish between individuals who will and will not have events and to reclassify individuals into clinically actionable risk strata. However, previous systematic reviews were conducted before the release of newly recommended risk tools and treatment thresholds. Additionally, these reviews did not address the ability of nontraditional risk factors to improve on existing models’ overestimation or underestimation of risk.

Based on previous systematic reviews, the US Preventive Services Task Force (USPSTF) issued recommendations in 2009 and 2013, concluding that the evidence was insufficient to assess the benefits and harms of using nontraditional risk factors for cardiovascular risk assessment (I statements). This review sought to identify and appraise updated evidence on the benefits and harms of nontraditional risk factor assessment and therapy guided by the ABI, hsCRP level, and CAC score to support the USPSTF in updating its 2009 and 2013 recommendations.

Methods

Scope of Review
This review addressed 5 key questions (KQs) as shown in Figure 1. Methodological details including study selection, a list of excluded studies, data analysis methods, and additional subpopulation results, as well as detailed descriptions of all models, are available in the evidence reviews for the USPSTF. The questions are depicted by linkages that relate interventions and outcomes. Further details are available in the USPSTF procedure manual. 36 ABI indicates ankle-brachial index; CAC, coronary artery calcium; CVD, cardiovascular disease; hsCRP, high-sensitivity C-reactive protein; MI, myocardial infarction.

Key questions

1. Compared with the Pooled Cohort Equations or Framingham risk factors alone, does risk assessment of asymptomatic adults using nontraditional risk factors—followed by treatment specific to risk level—lead to reduced incidence of cardiovascular events (eg, myocardial infarction, cerebrovascular accident), mortality, or both?

2. Does use of nontraditional risk factors in addition to traditional risk factors to predict cardiovascular disease risk improve measures of calibration, discrimination, and risk reclassification?

3. What are the harms of nontraditional risk factor assessment?

4. Does treatment guided by nontraditional risk factors, in addition to traditional risk factors, lead to reduced incidence of cardiovascular events (eg, myocardial infarction, cerebrovascular accident), mortality, or both?

5. What are the harms of treatment guided by nontraditional risk factors?

Evidence reviews for the US Preventive Services Task Force (USPSTF) use an analytic framework to visually display the key questions that the review will address to allow the USPSTF to evaluate the effectiveness and safety of a preventive service. The questions are depicted by linkages that relate interventions and outcomes. Further details are available in the USPSTF procedure manual. 36 ABI indicates ankle-brachial index; CAC, coronary artery calcium; CVD, cardiovascular disease; hsCRP, high-sensitivity C-reactive protein; MI, myocardial infarction.

* Risk factors: age, sex, blood pressure, levels of total and high-density lipoprotein cholesterol, smoking, diabetes, race/ethnicity.
In total, the current review included 54 articles (43 studies); studies may appear in more than 1 key question (KQ). Reasons for exclusion: Aim: Study aim was not relevant. Setting: Study was not conducted in a country relevant to US practice, or not conducted in, recruited from, or feasible for primary care or a health care system. Population: Study was not conducted in adults without known cardiovascular disease. Outcomes: Study did not report required outcomes. Intervention: Intervention was out of scope. Study design: Study did not use an included design. Quality: Study was poor quality. Base model: Eligible base models had to include age, sex, systolic blood pressure, antihypertensive medication use, total cholesterol, high-density lipoprotein cholesterol, and current smoking status; eligible base models could not include additional risk factors. Comparator: Study did not include an eligible model for comparison. Ancillary: The article met inclusion criteria, but more complete data such as a larger sample or longer follow-up was abstracted from a different included article.  

Data Sources and Searches 
MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials were searched through May 22, 2017, to identify literature published after 2 reviews for the USPSTF (eMethods in the Supplement). All studies in the prior reviews were also evaluated, as well as reference lists of other systematic reviews. ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform were searched for relevant ongoing trials. Since May 22, 2017, ongoing surveillance through article alerts and targeted searches of journals with a high impact factor and journals relevant to the topic was conducted to identify major studies published in the interim that may affect the conclusions or understanding of the evidence and therefore the related USPSTF recommendation. The last surveillance was conducted on February 7, 2018, and identified no additional relevant studies.

Study Selection 
Investigators reviewed 22,707 unique citations and 483 full-text articles against a priori inclusion criteria (Figure 2; eTable 1 in the Supplement).

Studies of adults without known CVD that were published in English were eligible for inclusion. For each KQ, subpopulation analyses by sex, race/ethnicity, and diabetes were identified a priori. For KQ1, trials comparing traditional risk assessment with traditional risk assessment plus the ABI, hsCRP level, or CAC score that reported patient health outcomes (ie, CVD events, mortality, or both) were included. For KQ2, individual participant data meta-analyses, trials, and well-designed prospective cohort studies evaluating risk prediction in models with traditional risk factors (base model), compared with models additionally including the ABI, hsCRP level, or CAC score (extended model), were included. KQ2 studies were required to include a measure of calibration, discrimination, or reclassification (Box). Eligible base models were the Pooled Cohort Equations, the Framingham Risk Score, or models including the same variables: age, sex, systolic blood pressure, use of antihypertensive medication, total cholesterol level, high-density lipoprotein cholesterol level, and smoking status. Models were eligible with or without inclusion of race/ethnicity and diabetes status. Studies with additional variables in their base models were excluded, as this would have precluded isolation of the effect of the nontraditional risk factor of interest. Extended models that incorporated multiple nontraditional risk factors were excluded if the effect of newly added risk factors could not be isolated.
Box. Definitions of Risk Prediction Measures Included for KQ2

Calibration refers to the agreement between observed and predicted outcomes. Calibration plots and observed to expected ratios were identified a priori as preferable measures of calibration because of their ability to indicate direction of miscalibration, capacity to compare across models, and intuitive graphic interpretation.

Discrimination is the ability to distinguish between individuals who will and will not have an event. This analysis uses the change in C statistics to measure discrimination. The C statistic is the probability that, for a randomly selected pair of individuals, 1 with disease and the other without, the person with disease will have the higher estimated disease probability according to the model.25

Reclassification reflects the ability of a new model to appropriately reassign people into different risk strata. The most commonly used measure in this review was the Net Reclassification Index (NRI), which captures the reclassification of individuals that occurs when traditional risk assessment is enriched with the ABI, hsCRP level, or CAC score, where clinically meaningful risk strata are typically defined by treatment thresholds. Reclassification to a higher risk category is considered appropriate when the individual has an outcome, and reclassification to a lower risk category is considered appropriate when an individual does not have an outcome. The NRI is the sum of differences in proportions of individuals moving up a risk category minus those moving down a risk category with a cardiovascular disease outcome (known as the event NRI), plus the proportion moving down a risk category minus those moving up a risk category without an outcome (known as the nonevent NRI).26

Abbreviations: ABI, ankle-brachial index; CAC, coronary artery calcium; hsCRP, high-sensitivity C-reactive protein; KQ, key question.

For KQ3 and KQ5, trials, prospective and retrospective cohort studies, and well-designed case-control studies examining harms of risk assessment or treatment guided by risk assessment were included. Harms included any serious adverse event requiring unexpected or unwanted medical attention resulting from risk assessment or harms from risk factor modification. For assessment of CAC score, radiation exposure from computed tomography (CT) was included as a potential harm. For KQ4, trials of treatment guided by nontraditional risk factor assessment in addition to traditional risk assessment vs no treatment or usual care that reported patient health outcomes (ie, CVD events, mortality, or both) were included.

Data Extraction and Quality Assessment

Two independent reviewers critically appraised the included studies using predefined criteria, with a third reviewer resolving disagreements (eTable 2 in the Supplement). Articles were rated as good, fair, or poor quality. In general, a good-quality study met all criteria. A fair-quality study did not meet, or it was unclear whether it met, at least 1 criterion but had no known important limitations that could invalidate its results. A poor-quality study had a single fatal flaw or multiple important limitations. Poor-quality studies were excluded from the review. Two studies, both for KQ2, were excluded as poor quality because of nonrepresentative sampling of patients, self-reported outcomes, limited duration of follow-up, and/or a small number of CVD events. One reviewer abstracted descriptive and outcome data from each included study into standardized evidence tables; a second checked for accuracy and completeness.

Data Synthesis and Analysis

Data for each KQ are summarized narratively according to nontraditional risk factor. Quantitative analyses were not conducted because of the limited number of studies for each KQ or methodological and clinical heterogeneity, including differences in outcomes and treatments evaluated. KQ2 results are stratified by calibration, discrimination, or reclassification (Box). A full description of selected measures for each domain and their limitations is available in eTable 3 in the Supplement.

Because there is no guidance in existing literature about how to characterize the magnitude or clinical meaning of changes in discrimination (KQ2), the following definitions were used for practical reasons. For changes in the C statistic, the term “large” is used to denote changes of 0.1 or greater, “moderate” for changes of 0.05 to 0.1, “small” for changes of 0.025 to 0.05, and “very small” for changes less than 0.025. C statistics range from 0.5 to 1.0; the 0.1 cutpoint for “large” was set because it represents 20% of the possible range. A change in C statistic of 0.025 approximates a 5% higher sensitivity when specificity is 50%. KQ2 analyses were additionally stratified by (1) type of model design—ie, published coefficients vs model development; (2) choice of the Pooled Cohort Equations or the Framingham Risk Score as the base model; and (3) prediction of global CVD outcomes, which include coronary and noncoronary events. Published coefficient model studies are risk prediction studies evaluating the added prognostic value of the ABI, hsCRP level, or CAC score that preserved the coefficients from the published Pooled Cohort Equations or Framingham Risk Score, which are readily available for use in clinical practice. This is in contrast to model development studies, which fit entirely new models with locally developed coefficients, such that these models are generally not available in the public domain for clinical use. Studies that used the original published coefficients of the Framingham Risk Score or the Pooled Cohort Equations were considered preferable to model development studies because of their applicability.

To help illustrate the clinical meaning of the Net Reclassification Index (NRI), event NRI, and nonevent NRI for each nontraditional risk factor, reclassification tables from selected studies were modified to correspond to current treatment thresholds and show reclassification in both absolute and relative terms. A 2016 Multi-Ethnic Study of Atherosclerosis (MESA) published coefficient analysis (which evaluated all 3 nontraditional risk factors using both Pooled Cohort Equations and Framingham Risk Score base models) and 2 individual patient data meta-analyses (available for the ABI and hsCRP level) were selected based on the applicability of the models evaluated, the ability to compare across risk factors, the larger populations represented therein, and the reporting of reclassification tables allowing for this analysis. For the Framingham Risk Score–based analyses, which used 3 risk strata, the top 2 risk strata (10%–20% and >20%) were combined to conform to the current USPSTF recommendation to initiate preventive low-dose aspirin and statins based on a 10-year CVD risk of 10% or greater.
Table 1. Evidence Landscape

<table>
<thead>
<tr>
<th>KQ</th>
<th>Any Outcome</th>
<th>Calibration</th>
<th>Discrimination</th>
<th>Reclassification</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ3</td>
<td>ABI NA</td>
<td>10 articles</td>
<td>38,39,52,56,58,65,70,74,81,88</td>
<td>5 articles</td>
</tr>
<tr>
<td>KQ4</td>
<td>hsCRP NA</td>
<td>25 articles</td>
<td>24,25,38,40,46,47,49,54,58-60</td>
<td>9 articles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15 articles</td>
</tr>
<tr>
<td>KQ5</td>
<td>CAC 1 RCT</td>
<td>2 RCTs</td>
<td>67,75</td>
<td>2 RCTs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19 articles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4 cohorts</td>
</tr>
</tbody>
</table>

**Abbreviations:** ABI, ankle-brachial index; CAC, coronary artery calcium; hsCRP, high-sensitivity C-reactive protein.

Result

In total, 43 studies reported in 54 publications (n = 267,244) were included (Figure 2). Thirty-eight unique trials or cohorts are represented in the included literature. For all KQs, additional descriptive and outcome data are available in the full report.

**Benefits of Risk Assessment**

**Key Question 1.** Compared with the Pooled Cohort Equations or Framingham risk factors alone, does risk assessment of asymptomatic adults using nontraditional risk factors lead to reduced incidence of cardiovascular events, mortality, or both?

Only 1 fair-quality trial evaluated the incremental value of nontraditional risk factor assessment on CVD events. The Early Identification of Subclinical Atherosclerosis by Non-invasive Imaging Research (EISNER) trial (n = 2137), conducted in the United States, randomized volunteers to undergo CT scanning for CAC scoring in addition to the Framingham Risk Score vs no CAC scoring before risk factor counseling. Participants were middle-aged adults with CVD risk factors but no known CVD or symptoms. This study found no statistically significant difference in myocardial infarction, mortality, or combined myocardial infarction and mortality at 4 years between the 2 groups; however, the trial did not have adequate sample size and length of follow-up to detect differences in these outcomes, as the primary outcome of this trial was a change in CVD risk factors and Framingham Risk Score at 4 years.

**Performance Characteristics of Risk Assessment Measures**

**Key Question 2.** Does use of nontraditional risk factors in addition to traditional risk factors to predict cardiovascular disease risk improve measures of calibration, discrimination, and risk reclassification?

Thirty-three studies reported in 43 articles evaluated the ABI, hsCRP level, and/or CAC score in addition to traditional cardiovascular risk assessment and reported 1 or more measures of calibration, discrimination, and/or risk reclassification. Ten studies evaluated the ABI, 25 studies evaluated hsCRP level, and 19 studies evaluated CAC score (Table 1). The evidence base for the ABI and hsCRP included large individual patient data meta-analyses? in these cases, the meta-analyses are discussed as the central piece of evidence and other studies of individual cohorts are discussed in relation to these meta-analyses.

In general, participants included in studies represented a broad range of primary prevention populations to whom CVD risk assessment is applicable. Most individuals were recruited as part of population-based cohort studies; additional populations were from cohorts derived from treatment trials and 1 registry of individuals who received CAC scores. Some analyses excluded people with known diabetes or already taking statins. Most included studies have important limitations in their applicability to current clinical practice. Fourteen studies (32.6%) used the published coefficients of publicly available Framingham Risk Score or Pooled Cohort Equations models. Moreover, the "low," "intermediate," and "high" risk strata used in many studies are not consistent with those used in current practice, where a
single threshold, such as 7.5% or 10% 10-year risk, is used for decision making to initiate preventive therapy.

Calibration
Limited data are available to inform whether the addition of the ABI (5 studies), hsCRP level (9 studies), or CAC score (8 studies) can improve agreement between predicted and observed events of risk assessment models. Across all 3 nontraditional risk factors, less than half of studies reported a measure of calibration (Table 1). Of these, only 24,25 reported the preferred measures of calibration plots and observed to expected ratios, and these are reported only for studies evaluating hsCRP level. These studies showed that the addition of hsCRP can improve calibration in some risk groups but worsen it in others; however, the small number of events—particularly in lower-risk groups—precludes definitive conclusions.

Discrimination
Ankle-Brachial Index | A large body of literature (10 studies; n = 79 583) consistently shows that the addition of the ABI to risk prediction models generally results in no to small improvement in a model’s ability to distinguish between individuals who will and will not have cardiovascular events (C statistic change of ~0.006 to 0.0316); however, improvements in discrimination can be large (eg, change of 0.112) when base models perform poorly (eTable 4 in the Supplement). Most models are not directly applicable to current practice because most are model development studies rather than published coefficient analyses, and evidence for the addition of the ABI to the Pooled Cohort Equations is sparse. The 1 study evaluating the Pooled Cohort Equations, a 2016 analysis of the MESA study38 showed moderate base model discrimination of 0.74 and no statistically significant change in discrimination of 0.01 when the ABI was added. The central piece of evidence for the addition of the ABI to the Framingham Risk Score is the Emerging Risk Factors Collaboration individual patient data meta-analysis, which involved 166 596 participants and 13 568 hard CVD events; this was a model development study with a moderate base model discrimination of 0.714.40 This analysis showed that the addition of hsCRP level increased discrimination by 0.004 (95% CI, 0.0028 to 0.0050) for predicting a composite of fatal and nonfatal myocardial infarction and cerebrovascular accident. Exploratory subgroup analyses showed a very small, statistically significant improvement in men and no change in women; the P value for heterogeneity was less than .001. Results were inconsistent for studies not included in the individual patient data meta-analysis, ranging from a worsening of discrimination with the addition of hsCRP level24 to a small improvement59,71; however, estimates of even small improvement likely represent an upper bound because of study design limitations.

CAC Score | The evidence (18 studies; n = 60 486) shows that the addition of CAC score to traditional risk assessment models results in the largest improvement in discrimination of all nontraditional risk factors evaluated, with change in the C statistic ranging from 0.018 to 0.144 (eTable 6 in the Supplement); however, there is no available individual patient data meta-analysis allowing for robust sex-stratified analyses. Two studies evaluated published coefficient Pooled Cohort Equations models; however, both were analyses of the same cohort.38,51 The MESA analysis by Yeboah et al,38 which had moderate base model discrimination of 0.74, found a very small to small, statistically significant improvement of 0.02 with the addition of CAC score to the model. The MESA analysis by Fudim et al,51 which presented only sex-stratified results, demonstrated similar findings, although statistically significant only for men. Results for published coefficient Framingham Risk Score base models were statistically significant and showed a slightly higher magnitude of change. For example, the 2016 MESA analysis found a statistically significant improvement in discrimination of 0.04.38 Results for model development studies were generally consistent with results from published coefficient analyses.

Recalssification
Ankle-Brachial Index | A large but heterogeneous body of literature (9 studies; n = 46 979) showed that the addition of the ABI to risk prediction models can improve the appropriate reclassification of individuals into clinically meaningful risk strata; NRIs are at best less than 0.1 and are usually much smaller and often nonsignificant (Table 2). There was considerable variation in the definitions of risk strata used across studies, and only 1 study used the Pooled Cohort Equations, with a treatment threshold of 7.5% 10-year risk.38 In this study, the total NRI was 0.017 (95% CI, −0.024 to 0.058). The individual patient data meta-analysis published coefficient Framingham Risk Score analysis used 10-year risk strata of less than 10%, 10% to 19%, and 20% or greater.39 The NRIs based on these 3 strata are no longer as relevant because current treatment thresholds to
<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>NRI Range (95% CI)</th>
<th>Overall</th>
<th>Event</th>
<th>Nonevent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pooled Cohort Equations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABI (1 study, n = 5185)</td>
<td>0.017 (−0.031 to 0.058)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.013 (−0.034 to 0.051)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.004 (−0.004 to 0.011)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>hsCRP (1 study, n = 5185)</td>
<td>0.024 (−0.015 to 0.067)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.028 (−0.013 to 0.077)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.005 (−0.015 to 0.003)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>CAC (2 studies, n = 11 927)</td>
<td>0.119 (0.080 to 0.256)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.178 (0.080 to 0.256)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.059 (−0.075 to 0.030)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td><strong>Framingham Risk Score</strong>&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABI (7 studies, n = 30 373)</td>
<td>0.001 (−0.06 to 0.058) to 0.079 (NR)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>−0.01 (−0.07 to 0.045) to 0.041 (−0.01 to 0.108)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>−0.003 (−0.008 to 0.004) to 0.027 (NR)&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>IPD meta-analysis (1 study, n = 11 421)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>0.096 (0.061 to 0.164)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.145 (0.101 to 0.189)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>−0.051 (−0.059 to −0.043)&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>0.043 (0.008 to 0.076)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.026 (−0.005 to 0.058)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.016 (0.004 to 0.027)&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>hsCRP (13 studies, n = 37 742)</td>
<td>0.003 (−0.028 to 0.026) to 0.120 (NR)</td>
<td>0.005 (−0.027 to 0.027) to 0.1091 (0.022 to 0.196)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>−0.012 (−0.029 to 0.006)&lt;sup&gt;h&lt;/sup&gt; to 0.099 (NR)</td>
<td></td>
</tr>
<tr>
<td>IPD meta-analysis (1 study, n = 72 574)</td>
<td>0.0152 (0.0078 to 0.0227)</td>
<td>0.0146 (0.0073 to 0.0219)</td>
<td>0.0006 (−0.0009 to 0.0022)</td>
<td></td>
</tr>
<tr>
<td>CAC (13 studies, n = 46 362)</td>
<td>0.084 (0.024 to 0.196) to 0.35 (0.11 to 0.58)</td>
<td>0.119 (0.045 to 0.239) to 0.333 (0.156 to 0.511)</td>
<td>−0.041 (−0.058 to −0.024) to 0.06 (0.028 to 0.092)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ABI, ankle-brachial index; CAC, coronary artery calcium; hsCRP, high-sensitivity C-reactive protein; IPD, individual patient data; NR, not reported; NRI, net reclassification index.

a Categorical NRIs are measures of reclassification that are influenced by event rates and risk cutoffs; continuous NRIs define upward and downward movement by a change in predicted probabilities. Because these measures are not comparable, only categorical NRIs, which are the most commonly reported, are included in this table.

b Two-category risk strata (Yeboah, 2016<sup>38</sup>): 7.5% or greater, less than 7.5%.

c Evidence from published coefficient studies only; no model development studies available.

d Three-category risk strata: varies.

e Range of −0.011 (not reported; not significant) to 0.159 (not reported; not significant) when sex-stratified analyses considered; the upper and lower ranges are for men and women, respectively, from the same study.<sup>30</sup>

f Many confidence intervals not reported or not calculable.

h Confidence intervals were calculated using simple variance method.
initiate preventive therapies have been lowered to 7.5% or 10%, and reported NRI will capture movement between the middle and upper categories. Consistent with findings for discrimination, the addition of the ABI showed a larger improvement in women than in men. In contrast to the findings for the Pooled Cohort Equations, this improvement was statistically significant for the Framingham Risk Score base model: 0.096 (95% CI, 0.061 to 0.164) for women and 0.043 (95% CI, 0.008 to 0.076) for men. Event NRIs (the appropriate upward reclassification of individuals having CVD events) were larger than nonevent NRIs (the appropriate downward reclassification of individuals not having CVD events). However, the nonevent NRI for women was negative and statistically significant, meaning that more women without events were inappropriately reclassified to a higher risk stratum than appropriately reclassified to a lower risk stratum. Consideration of the event and nonevent NRI separately is important because the total NRI does not weight for the proportion of individuals who have or do not have events; in the included primary prevention populations, it is much more common not to have an event. Cohorts not included in the individual patient data meta-analysis showed a similar magnitude of reclassification with the addition of the ABI, with mixed statistical significance.

hsCRP Level | The body of evidence for the addition of hsCRP level to risk prediction models (15 studies; n = 115,686) shows inconsistent evidence for an improvement in the appropriate reclassification of individuals into clinically meaningful risk strata; evidence from the largest analysis suggests an overall NRI of no greater than 0.02 (Table 2). As with the evidence for the ABI, risk strata were defined variably among studies, with few using thresholds applicable to current practice. Additionally, many studies used 3 risk strata, which is no longer as relevant to current clinical practice. Only 1 study used the Pooled Cohort Equations with treatment threshold of 7.5% 10-year risk, and this was the same study that evaluated the addition of the ABI to the Pooled Cohort Equations.38 In this study, the total NRI was 0.024 (95% CI, −0.015 to 0.067). The IPD meta-analysis model development study (n = 72,574) used 10-year risk strata of less than 10%, 10% to 19%, and 20% or greater.40 The total NRI and event NRI were 0.0152 (95% CI, 0.0078 to 0.0227) and 0.0146 (95% CI, 0.0073 to 0.0219), respectively. However, the nonevent NRI was much smaller and nonsignificant (0.0006 [95% CI, −0.0009 to 0.0022]); about 92% of individuals in the individual participant data meta-analysis did not have events. Similar to results for discrimination, exploratory subgroup analyses by sex show that reclassification was greater for men than for women, but reclassification results were not significant for either sex individually. Two other studies offer some confirmatory evidence about a larger effect for men, although reporting inconsistencies preclude definitive conclusions.60,76 The results of cohorts not included in the individual patient data meta-analysis were also inconsistent with respect to magnitude and statistical significance. Predicted outcome, definitions of risk strata, and case mix did not explain differences among studies; however, these comparisons were limited by several concurrent sources of heterogeneity.

CAC Score | Of the 3 nontraditional risk factors evaluated in this review, the addition of CAC score to traditional risk assessment models has the strongest effect on appropriate reclassification (15 studies; n = 58,289). Results range from 0.084 (95% CI, 0.024 to 0.196) to 0.35 (95% CI, 0.11 to 0.58) and are consistently statistically significant; however, the nonevent NRI is negative in several studies (Table 2). In the absence of an individual patient data meta-analysis for CAC score, few reliable data are available to inform whether effect modification exists by sex. As for other nontraditional risk factors, studies used a variety of definitions of risk strata, which precludes definitive comparisons. The most applicable evidence is from the published coefficient Pooled Cohort Equations analysis,38 which reported an overall NRI of 0.119 (95% CI, 0.080 to 0.256). Numerically, the overall NRI was driven by the event NRI of 0.178 (0.080 to 0.256). The nonevent NRI was negative (−0.059 [95% CI, −0.075 to 0.030]), meaning that individuals not having events were inappropriately reclassified upward into higher risk and above a treatment threshold. While not statistically significant in this study, the negative nonevent NRI is statistically significant in other studies.55,58,68

Clinical Interpretation of Reclassification Across Nontraditional Risk Factors

The clinical significance of the NRI, event NRI, and nonevent NRI can be understood by illustrating the absolute number of people appropriately and inappropriately reclassified in terms of current treatment thresholds. Selected examples from included studies are provided in Table 3. In the example for the Pooled Cohort Equations, 76 people having CVD events were appropriately reclassified upward when CAC score was added to risk assessment, and 19 people who had events were inappropriately reclassified downward who had a CVD event—a net improvement of 57 individuals among the 320 having events, about 18 events per 100 people (reported event NRI of 0.178 [95% CI, 0.080 to 0.256]). However, in the primary prevention populations to which CVD risk assessment with the Pooled Cohort Equations or Framingham Risk Score applies, the majority of people will not experience a CVD event. With the addition of CAC score to risk assessment, 202 people not having events are appropriately reclassified downward, but 496 people are inappropriately reclassified upward on net, a worsening of reclassification of 294 individuals of 4865 not having events—or about 6 events per 100 people (reported nonevent NRI of −0.059 [95% CI, −0.075 to 0.030]). Therefore, the NRI of 0.119 [95% CI, 0.080 to 0.256]) does not convey that for CAC score a sizeable proportion of individuals who are not having events will now be considered for treatment. The addition of the ABI to the Framingham Risk Score in women showed a similar pattern.

Harms of Risk Assessment

Key Question 3. What are the harms of nontraditional risk factor assessment?

Evidence for the harms of nontraditional risk factor assessment was limited to 8 studies evaluating CAC score; no eligible studies evaluated the potential harms of the ABI or hsCRP level. Four studies reported radiation exposure for CT imaging to obtain CAC score,43,56,62,75 and 5 studies reported other potential adverse events from CAC score measurement such as psychological outcomes, adverse cardiovascular events, and health care utilization.43,56,66,67,75,80

Overall, the radiation exposure or effective radiation dose per CT examination is low: 2 mSv or less. Based on 2 studies (n = 1619)—a small randomized clinical trial (RCT) and a subsample

Downloaded From: https://jamanetwork.com/ by a Non-Human Traffic (NHT) User on 04/26/2019
from a population-based cohort—risk assessment with CAC score does not appear to cause short-term mental distress.66,67 Additionally, 2 studies (n = 11 364) using administrative claims data showed that risk assessment with CAC score did not appear to paradoxically increase CVD events.45,80 However, a small body of evidence with study design and applicability limitations shows mixed findings for the effect of CAC score on downstream health care utilization.45,75,80 Both the EISNER RCT and 1 cohort using administrative data found no statistically significant increase in cardiac imaging or revascularization procedures in patients receiving vs not receiving CAC screening in follow-up of 4 years or 6 months, respectively.45,75 However, 1 study using Medicare claims data found a greater number of subsequent cardiac imaging tests and revascularization in asymptomatic people who received CAC screening compared with those receiving hsCRP or lipid screening.80

**Benefits and Harms of Risk Factor–Guided Treatment**

**Key Question 4.** Does treatment guided by nontraditional risk factors, in addition to traditional risk factors, lead to reduced incidence of cardiovascular events, mortality, or both?

No trials compared treatment guided by nontraditional risk factors in addition to traditional risk factors vs no treatment or usual care. In the absence of this evidence, studies in which preventive therapies were guided by nontraditional risk factors alone, without formal traditional risk assessment, were included. Four such RCTs reported the outcomes of CVD events, mortality, or both.41,42,50,64,73 Two of these trials evaluated aspirin in individuals with an abnormal ABI; 1 trial evaluated high-intensity statins in those with an abnormal hsCRP level, and 1 trial evaluated moderate-intensity statins in those with an abnormal CAC score.

Two good-quality RCTs (n = 4626) in asymptomatic adults with an abnormal ABI—including 1 trial exclusively in participants with diabetes—did not find any statistically significant benefit for aspirin (100 mg daily) on reducing CVD outcomes or all-cause mortality compared with placebo after approximately 7 to 8 years of follow-up.42,50 However, neither trial used the conventional 0.90 threshold for an abnormal ABI; 1 defined an abnormal ABI as 0.95 or less and the other as 0.99 or less. One good-quality RCT (n = 17 802) in asymptomatic people with elevated hsCRP level (≥2.0 mg/L [19.05 nmol/L]) but normal low-density lipoprotein cholesterol level (<130 mg/dL [3.37 mmol/L]) found a significant relative reduction in CVD events for rosuvastatin (20 mg daily) compared with placebo (hazard ratio, 0.56 [95% CI, 0.46 to 0.69]) at approximately 2 years.73 One fair-quality trial (n = 1005) in asymptomatic people with elevated CAC score, defined as 80th percentile or greater for age and sex, and with low-density lipoprotein cholesterol levels less than 175 mg/dL (4.5 mmol/L), did not find any statistically significant benefit for atorvastatin (20 mg daily) on reducing CVD outcomes compared with placebo after a mean of 4.3 years of follow-up; however, that study had a lower than expected number of events and was terminated early because of futility.

**Key Question 5.** What are the harms of treatment guided by nontraditional risk factors?

Three of the 4 RCTs included for KQ4 reported harms of aspirin or statins guided by nontraditional risk factor assessment.42,50,73 No other studies evaluating harms met inclusion criteria. Neither aspirin trial (n = 4626) found a statistically significant increase in hemorrhagic cerebrovascular accident after approximately 7 to 8 years of follow-up, although analyses are limited by a rare event rate.42,50 In the 1 trial reporting major bleeding, the association with low-dose aspirin approached statistical significance (hazard ratio, 1.71 [95% CI, 0.99 to 2.97]).50 The trial evaluating high-intensity statins in adults with elevated hsCRP levels (n = 17 802) found evidence of an increased incidence of diabetes in the treatment group

### Table 3. Selected Examples of Appropriate and Inappropriate Reclassification Using Current Treatment Thresholds (≥7.5% for the Pooled Cohort Equations, ≥10% for Framingham Risk Score)

<table>
<thead>
<tr>
<th>Nontraditional Risk Factor</th>
<th>People With CVD or CHD Event</th>
<th>People Without CVD or CHD Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute No. of People Reclassified</td>
<td>Appropriately Reclassified</td>
<td>Inappropriately Reclassified</td>
</tr>
<tr>
<td>PCE (Predicted Outcome: Hard CVD)80</td>
<td>ABI</td>
<td>hsCRP</td>
</tr>
<tr>
<td>Absolute No. of People Reclassified</td>
<td>Appropriately Reclassified</td>
<td>Inappropriately Reclassified</td>
</tr>
<tr>
<td>ABI</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>hsCRP</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>CAC</td>
<td>76</td>
<td>19</td>
</tr>
<tr>
<td>FRS (Predicted Outcome: Hard CHD)45</td>
<td>ABI</td>
<td>hsCRP</td>
</tr>
<tr>
<td>Absolute No. of People Reclassified</td>
<td>Appropriately Reclassified</td>
<td>Inappropriately Reclassified</td>
</tr>
<tr>
<td>ABI</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>IPD meta-analysis</td>
<td>Men6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Women6</td>
<td>46</td>
</tr>
<tr>
<td>hsCRP</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IPD meta-analysis</td>
<td>162</td>
<td>131</td>
</tr>
<tr>
<td>CAC</td>
<td>26</td>
<td>5</td>
</tr>
</tbody>
</table>

Abbreviations: ABI, ankle-brachial index; CAC, coronary artery calcium; CHD, coronary heart disease; CVD, cardiovascular disease; FRS, Framingham Risk Score; hsCRP, high-sensitivity C-reactive protein; IPD, individual participant data.  
* Rounded to whole numbers.  
+ MESA (Multi-Ethnic Study of Atherosclerosis) cohort from Yeboah, 201638 (N = 5185 in reclassification analyses).  
* IPD meta-analysis study for the ABI reported in ABI Collaboration, 201479 (N = 6459 women and N = 4962 men in reclassification analyses).  
* IPD meta-analysis study for hsCRP reported in Emerging Risk Factors Collaboration, 201245 (N = 72 574 in reclassification analyses).
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies*</th>
<th>Summary of Findings by Outcome</th>
<th>Consistency</th>
<th>Reporting Bias</th>
<th>Study Quality</th>
<th>Body of Evidence Limitations</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ1: Direct Evidence for Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVD events or mortality</td>
<td>1 study, n = 2137 (RCT)</td>
<td>No statistically significant difference in MI and/or mortality at 4 y between those who received CAC screening vs those who did not</td>
<td>Consistency NA</td>
<td>Precision NA</td>
<td>Fair</td>
<td>Single trial; insufficient sample size and length of follow-up to detect differences in patient health outcomes</td>
<td>US-based trial, volunteer sample</td>
</tr>
<tr>
<td><strong>KQ2: ABI Risk Prediction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration</td>
<td>5 studies, n = 26,286 (1 IPD meta-analysis, 4 cohort)</td>
<td>Based on 1 IPD meta-analysis, various measures demonstrate that the addition of ABI to FRS can improve model fit. However, the clinical meaning of changes in these measures of calibration is unclear.</td>
<td>Reasonably consistent</td>
<td>Undetected</td>
<td>1 Good, 4 fair</td>
<td>No preferred measures of calibration</td>
<td>No evidence for PCE. IPD meta-analysis in white patients only</td>
</tr>
<tr>
<td>Discrimination</td>
<td>10 studies, n = 79,583 (1 IPD meta-analysis, 8 cohort)</td>
<td>Based on 1 IPD meta-analysis, ABI can result in large improvement in discrimination when added to FRS in women, but not men, primarily because of poorer discrimination of the base model (using published coefficients) in women but not men. The incremental benefit in IPD meta-analysis model development analysis was very small for both men and women, owing to improved base model discrimination.</td>
<td>Reasonably consistent</td>
<td>Reasonably precise</td>
<td>Undetected</td>
<td>Adequate power for sex-stratified analyses limited to IPD meta-analysis; differences in study population, base models, and outcomes predicted limit direct comparison across studies</td>
<td>Limited evidence for PCE. IPD meta-analysis in white patients only</td>
</tr>
<tr>
<td>Risk reclassification</td>
<td>9 studies, n = 46,979 (1 IPD meta-analysis, 7 cohort)</td>
<td>Based on 1 IPD meta-analysis, ABI can result in improvement in reclassification when added to FRS (using published coefficient) in women, but not men; most promising for women at intermediate risk for hard CHD events. However, examination of separate components of the NRI (event and nonevent NRI) suggests that improvement in reclassification comes from women who had events being appropriately reclassified as having a higher risk, in contrast, women who did not have a cardiovascular event (the majority of the population) were inappropriate reclassified as having a higher risk (ie, a negative nonevent NRI). Improvement in NRI was not observed in the model development IPD meta-analysis.</td>
<td>Reasonably consistent</td>
<td>Reasonably precise</td>
<td>Undetected</td>
<td>Adequate power for sex-stratified analyses limited to IPD meta-analysis; differences in study population, base models, and outcomes predicted limit direct comparison across studies NRI not weighted for prevalence of events and nonevents, so reclassification benefit may be overstated</td>
<td>Limited evidence for PCE. IPD meta-analysis in white patients only Risk categories on which NRI analyses are based can vary across studies and may not apply to current practice</td>
</tr>
<tr>
<td><strong>KQ2: hsCRP Risk Prediction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration</td>
<td>9 studies, n = 50,343 (8 cohort, 1 nested case-control)</td>
<td>Various measures demonstrate that the addition of hsCRP to traditional risk factors can improve model fit. However, the clinical meaning of changes in these measures is unclear. In model development studies, calibration plots suggested that the addition of hsCRP can improve model fit in some but not all risk groups.</td>
<td>Reasonably consistent</td>
<td>Undetected</td>
<td>2 Good, 7 fair</td>
<td>No preferred measures for most studies (and none for published coefficient models); no calibration statistics for the IPD meta-analysis</td>
<td>Limited evidence for PCE. Model development IPD meta-analysis only (calibration in model development less applicable to clinical practice)</td>
</tr>
</tbody>
</table>

(continued)
Table 4. Summary of Evidence (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies(^a)</th>
<th>Summary of Findings by Outcome</th>
<th>Consistency and Precision</th>
<th>Reporting Bias</th>
<th>Study Quality</th>
<th>Body of Evidence Limitations</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discrimination</strong></td>
<td>25 studies, (n = 265,704) (1 IPD meta-analysis, 18 cohort, 3 nested case-control, 1 case-cohort)</td>
<td>At best, improvement in discrimination from the addition of hsCRP to traditional cardiovascular risk assessment is small and more likely to occur in the context of a poorly discriminating base model. IPD meta-analysis model development study found very small improvement in discrimination from the addition of hsCRP to FRS to predict hard CHD</td>
<td>Inconsistent, Reasonably precise for IPD meta-analysis</td>
<td>Undetected</td>
<td>2 Good, 23 fair</td>
<td>Limited reporting of confidence intervals and statistical significance; differences in study population, base models, and outcomes predicted limit direct comparison across studies</td>
<td>Limited evidence for PCE. Model development IPD meta-analysis only (changes in discrimination in model development may be less applicable to clinical practice)</td>
</tr>
<tr>
<td>Risk reclassification</td>
<td>15 studies, (n = 115,686) (1 IPD meta-analysis, 13 cohort, 1 nested case-control)</td>
<td>NRIs from addition of hsCRP to FRS are inconsistent; 1 published coefficient PCE-based study suggested no improvement in reclassification. Best evidence from IPD meta-analysis showed a statistically significant NRI of 0.0152 (95% CI, 0.0078-0.0227). Sex-stratified analyses suggest that reclassification occurs in men but not women. The bias-corrected intermediate-risk group NRI from the IPD meta-analysis was still very small to small, at 0.027 (95% CI, 0.007-0.047).</td>
<td>Inconsistent, Reasonably precise for IPD meta-analysis</td>
<td>Undetected</td>
<td>2 Good, 13 fair</td>
<td>Comparisons across studies are limited by inconsistency in risk category definitions; sex-specific analyses reported rarely and more are needed to confirm the signal of effect modification by sex. Limited information on intermediate-risk group NRI, as analyses often underpowered and often cannot be corrected for bias.</td>
<td>Limited evidence for PCE. Risk categories on which NRI analyses are based can vary across studies and may not apply to current practice. Estimates of reclassification in model development may be less applicable to clinical practice.</td>
</tr>
<tr>
<td><strong>KQ2: CAC Risk Prediction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration</td>
<td>8 studies, (n = 29,775) (4 cohort(^b))</td>
<td>Limited model development studies using various measures demonstrate that the addition of CAC to traditional risk factors can improve model fit. However, the clinical meaning of changes in these measures is unclear.</td>
<td>Inconsistent, Unable to assess precision</td>
<td>Undetected</td>
<td>Fair</td>
<td>No preferred measures of calibration</td>
<td>No evidence for published coefficient models, calibration in model development less applicable to clinical practice, limited evidence in context of PCE</td>
</tr>
<tr>
<td>Discrimination</td>
<td>18 studies, (n = 60,486) (10 cohort(^c))</td>
<td>CAC in addition to traditional risk factor assessment results in changes of 0.018 to 0.144. Discrimination is not consistently greater in men or women.</td>
<td>Reasonably consistent, Reasonably precise</td>
<td>Undetected</td>
<td>Fair</td>
<td>Smaller cohorts compared with ABI and hsCRP body of evidence. No IPD meta-analysis limits understanding in differences by sex. Differences in study population, base models, and outcomes predicted limit direct comparison across studies.</td>
<td>Limited evidence in context of PCE. Non-population-based cohorts may not be broadly applicable.</td>
</tr>
<tr>
<td>Risk reclassification</td>
<td>15 studies, (n = 56,289) (9 cohort(^c))</td>
<td>CAC resulted in NRIs of 0.084 to 0.35 when added to traditional risk factor assessment. Evaluation of separate components of the NRI shows that improvements in NRI are consistently driven by event NRIs much larger than nonevent NRIs, which were commonly negative when reported and sometimes statistically significant. Reclassification is not consistently greater in men or women.</td>
<td>Reasonably consistent, Reasonably precise</td>
<td>Undetected</td>
<td>Fair</td>
<td>Smaller cohorts compared with ABI and hsCRP body of evidence. No IPD meta-analysis limits understanding in differences by sex. Limited information on intermediate-risk group NRI, as analyses often underpowered and often cannot be corrected for bias. Differences in study population, base models, and outcomes predicted limit direct comparison across studies. NRI not weighted for prevalence of events and nonevents, so the reclassification benefit may be overstated.</td>
<td>Non-population-based cohorts may not be broadly applicable. Risk categories on which NRI analyses are based can vary across studies and may not apply to current practice.</td>
</tr>
</tbody>
</table>
### Table 4. Summary of Evidence (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Summary of Findings by Outcome</th>
<th>Consistency and Precision</th>
<th>Reporting Bias</th>
<th>Study Quality</th>
<th>Body of Evidence Limitations</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ3: Harms of Screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation dose</td>
<td>4 studies, n = 11 473 (3 cohort, 1 RCT)</td>
<td>Effective dose of radiation per CT examination for screening CAC was low (≤2.1 mSv)</td>
<td>Reasonably consistent</td>
<td>Reasonably precise</td>
<td>Suspected</td>
<td>Fair</td>
<td>Only a limited subset of CAC studies included for KQ2 reported radiation dose. Dose not reported separately by EBCT vs MDCT.</td>
</tr>
<tr>
<td>Psychological outcomes</td>
<td>2 studies, n = 1619 (1 cohort, 1 RCT)</td>
<td>Screening CAC is not associated with subsequent depression, anxiety, or decline in overall mental health functioning up to 1 y</td>
<td>Reasonably consistent</td>
<td>Reasonably precise</td>
<td>Undetected</td>
<td>Fair</td>
<td>No studies for ABI or hsCRP. Only 1 study with a comparator group.</td>
</tr>
<tr>
<td>CVD events</td>
<td>2 studies, n = 11 364 (2 cohort)</td>
<td>No paradoxical increase in CVD events (MI, CVA, unstable angina) or all-cause mortality with screening CAC at approximately 1.5 y to 3 y of follow-up</td>
<td>Reasonably consistent</td>
<td>Reasonably precise</td>
<td>Undetected</td>
<td>Fair</td>
<td>No studies for ABI or hsCRP. Retrospective analyses of administrative data. Limited length of follow-up.</td>
</tr>
<tr>
<td>Health Care utilization</td>
<td>3 studies, n = 13 204 (2 cohort, 1 RCT)</td>
<td>Best-quality evidence from 1 RCT found no statistically significant increase in cardiac imaging or revascularization for screening CAC at 4 y of follow-up. Two retrospective cohort studies using differently assembled control groups had mixed findings. One study using Medicare claims data found a higher number of cardiac imaging and revascularization procedures associated with CAC as opposed to hsCRP or lipid screening.</td>
<td>Inconsistent</td>
<td>Imprecise</td>
<td>Undetected</td>
<td>Fair</td>
<td>No studies for ABI or hsCRP. No studies of downstream utilization due to incidental findings on CT for CAC. Two retrospective analyses of administrative data.</td>
</tr>
</tbody>
</table>

### KQ4: ABI-Guided Treatment Benefit

| CVD events | 2 studies, n = 4626 (2 RCT) | AAA and POPADAD found no benefit for ABI-guided low-dose aspirin (100 mg daily) in asymptomatic persons on composite CVD outcomes (MI, CVA, revascularization, or amputation) at approximately 7 y to 8 y of follow-up | Reasonably consistent | Reasonably precise | Undetected | Good | No ABI-guided statin trials. | Nontraditional threshold for ABI used in both trials. |

### KQ4: hsCRP-Guided Treatment Benefit

| CVD events | 1 study, n = 17 802 (RCT) | JUPITER found a benefit for hsCRP-guided, high-intensity statin (rosuvastatin [20 mg daily]) in asymptomatic persons on CVD outcomes at 1.9-y follow-up, 1.6% had a CVD event (MI, CVA, hospitalization for unstable angina, revascularization, or CVD mortality) in the statin group vs 2.8% in the placebo group (HR, 0.56 [95% CI, 0.46-0.69]). | Consistency NA | Reasonably precise | Undetected | Good | No hsCRP-guided aspirin treatment trials. Trial stopped early, which may overestimate findings of benefit. | Threshold for hsCRP was 2.0 mg/L (19.05 nmol/L). Diverse population. |

(continued)
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Summary of Findings by Outcome</th>
<th>Consistency and Precision</th>
<th>Reporting Bias</th>
<th>Study Quality</th>
<th>Body of Evidence Limitations</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ4: CAC-Guided Treatment Benefit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVD events</td>
<td>1 study, n = 1005 (RCT)</td>
<td>ST Francis Heart Study found no benefit for CAC-guided moderate-intensity statin (atorvastatin [20 mg daily]) in asymptomatic persons on composite CVD outcomes at approximately 4 y of follow-up</td>
<td>Consistency NA Imprecise</td>
<td>Undetected</td>
<td>Fair</td>
<td>No CAC-guided aspirin treatment trials</td>
<td>Trial not powered to detect a difference in outcomes</td>
</tr>
<tr>
<td><strong>KQ5: ABI-Guided Aspirin Treatment Harms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major bleeding</td>
<td>2 studies, n = 4626 (2 RCT)</td>
<td>AAA and POPADAD found no statistically significant difference in bleeding events between low-dose aspirin (100 mg daily) and placebo. However, AAA found a trend for increased bleeding events in the aspirin group (2.0%) vs placebo (1.2%) (HR, 1.71 [95% CI, 0.99-2.97]) at 8.2 y of follow-up</td>
<td>Reasonably consistent Imprecise</td>
<td>Undetected</td>
<td>Good</td>
<td>Limited follow-up</td>
<td>Likely not powered to detect a difference in bleeding events</td>
</tr>
<tr>
<td><strong>KQ5: hsCRP-Guided Statin Treatment Harms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious adverse events</td>
<td>1 study, n = 17 802 (RCT)</td>
<td>JUPITER found a statistically significant increase in incident diabetes but not in other serious adverse events. There were 3.0% cases of diabetes in the rosuvastatin group vs 2.4% cases in the placebo group (HR, 1.25 [95% CI, 1.04-1.50]) at 1.9 y of follow-up</td>
<td>Reasonably consistent Reasonably precise</td>
<td>Undetected</td>
<td>Good</td>
<td>Limited follow-up</td>
<td>This trial should be interpreted in the context of the larger body of evidence on adverse events from high-intensity statins.</td>
</tr>
</tbody>
</table>

Abbreviations: AAA, Aspirin for Asymptomatic Atherosclerosis; ABI, ankle-brachial index; CAC, coronary artery calcium; CHD, coronary heart disease; CT, computed tomography; CVA, cerebrovascular accident; CVD, cardiovascular disease; EBCT, electron beam computed tomography; FRS, Framingham Risk Score; HR, hazard ratio; hsCRP, high-sensitivity C-reactive protein; IPD, individual patient data; JUPITER, Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin; KQ, key question; MDCT, multidetector computed tomography; MI, myocardial infarction; NA, not applicable; NRI, Net Reclassification Index; PCE, Pooled Cohort Equations; POPADAD, Prevention of Progression of Arterial Disease and Diabetes; RCT, randomized clinical trial; RR, relative risk.

* ns approximated by using the largest N analyzed in each cohort.
Discussion

Direct evidence from adequately powered clinical trials evaluating the incremental effect of the ABI, hsCRP level, or CAC score to improve health outcomes is lacking (Table 4; eTable 7 in the Supplement). This remains unchanged from the 2009 and 2013 USPSTF recommendations, which concluded that the evidence was insufficient to assess the benefits and harms of using nontraditional risk factors for cardiovascular risk assessment (I statements).14,15

Advances in the evaluation of risk prediction literature, as well as new evidence since the previous recommendations, has allowed for more detailed analyses of risk prediction studies that offer indirect evidence for the use of nontraditional risk factors (Figure 1). A substantial body of evidence evaluates the ability of the ABI, hsCRP level, or CAC score to improve on the discrimination and reclassification of traditional risk assessment, but most of this evidence is not readily applicable to current clinical practice because it does not use published coefficients, rarely evaluates the addition of nontraditional risk factors in the context of the Pooled Cohort Equations, and seldom uses current risk thresholds for aiding treatment decisions. Regardless, CAC score shows the largest potential for improvement in discrimination and reclassification. Improvements for the ABI are larger when base models perform poorly, such as the Framingham Risk Score in women. Results for hsCRP level are inconsistent. Based on selected included studies, examination of the absolute number of people appropriately and inappropriately reclassified using current treatment thresholds (ie, 7.5% or 10% 10-year CVD risk) demonstrates that even with favorable NRI for CAC score and ABI, more people will be incorrectly reclassified to a higher risk category, thus needing treatment, vs correctly reclassified as needing or not needing treatment.

A 2015 analysis of data from the MESA study found inconsistent results. That analysis, while valid, was designed to answer a different question about risk stratification and was excluded from this review because of its different study design. Multiple methodologic differences between this analysis and the included MESA analysis (ie, differences in definitions of risk strata, differences in risk classification of individuals with diabetes, differences in comparisons of CAC score alone vs CAC score plus Pooled Cohort Equations, and recalibration of the Pooled Cohort Equations) prohibit direct comparisons between the 2 studies’ findings. Therefore, while CAC score appears to be the most informative nontraditional risk factor, there remains uncertainty around the overall clinical effect of the inappropriate upward reclassification of individuals and of resultant downstream testing, including from incidental findings. While experts have argued that large-scale clinical trials evaluating the effectiveness of CAC screening on patient health outcomes are not feasible,96 these studies may be necessary to address this remaining clinical uncertainty.

Despite the recent proliferation of risk prediction studies, few clarify whether the addition of the ABI, hsCRP level, or CAC score can improve the calibration of traditional risk assessment. The lack of reporting of preferred measures such as calibration plots and observed to expected ratios limits the ability to interpret the clinical meaning of improvements in calibration. Even when preferred measures are reported, confidence in their results is usually limited by small numbers of events, especially among low-risk groups. As treatment thresholds have decreased over time, the implications of calibration in lower-risk groups have become especially important. Historically, the performance of risk prediction models has been focused on discrimination,93 so the sparse reporting of calibration measures is not surprising and is consistent with the findings of other systematic reviews.94 The focus on discrimination is incomplete because the C statistic is a rank order statistic; therefore a model can discriminate well but still systematically underestimate or overestimate risk.95

Limitations

The review report has several limitations. The review focused on just 3 nontraditional risk factors—the ABI, hsCRP level, and CAC score. Many more have been identified in the literature; however, these risk factors were selected because other synthesized literature identified them as having the greatest clinical potential.21,27 Additionally, the predictive value of traditional risk factors such as levels of total or high-density lipoprotein cholesterol was taken as given, but some literature suggests that these, too, might be very small to small when assessed in terms of the C statistic.40 Given the large volume of studies included for KQ2, some explicit exclusions were made so as to focus on the most clinically relevant analyses. Additionally, comparisons across studies are difficult because of the heterogeneity of base models, model type, predicted outcome, and definitions of risk strata as well as limited reporting of confidence intervals, statistical significance, and separate reporting of event and nonevent NRIs. In addition, 1 of the risk assessment analysis stratifications was whether a study preserved published coefficients or developed a new model; because methods were often only briefly described, the categorization represents a best guess for many studies, and, in cases of uncertainty, the study was categorized as model development.

Conclusions

There are insufficient adequately powered clinical trials evaluating the incremental effect of the ABI, hsCRP level, or CAC score in risk assessment and initiation of preventive therapy. Furthermore, the clinical meaning of improvements in measures of calibration, discrimination, and reclassification risk prediction studies is uncertain.
Critical revision of the manuscript for important intellectual content: Lin, Evans, Johnson, Smith. Statistical analysis: Johnson, Redmond, Smith. Obtained funding: Lin. Administrative, technical, or material support: Evans, Coppola. Supervision: Lin, Evans.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: This research was funded under contract HHSAG20001500071, Task Order 2, from the Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services, under a contract to support the USPSTF.

Role of the Funder/Sponsor: Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight, reviewed the report to ensure that the analysis met methodological standards, and distributed the draft for peer review. Otherwise, AHRQ had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, and approval of the manuscript findings. The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

Additional Contributions: We gratefully acknowledge the following individuals for their contributions to this project: Todd Hannon, MLS, Janelle Guirguis-Blake, MD, and Katherine Essick, BS (Kaiser Permanente Center for Health Research); Robert Platt, PhD (McGill University); Justin Mills, MD, MPH (Agency for Healthcare Research and Quality [AHRQ]); Elisabeth Kato, MD, MRP (formerly at AHRQ); and current and former members of the US Preventive Services Task Force (USPSTF) who contributed to topic deliberations. USPSTF members, peer reviewers, and federal partner reviewers did not receive financial compensation for their contributions.

Additional Information: A draft version of this evidence report underwent external peer review from 5 content experts (Nancy Cook, ScD, Harvard T. H. Chan School of Public Health; Donald Lloyd-Jones, MD, Northwestern University Feinberg School of Medicine; Gerry Fowkes, MD, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh; Matthew Budoff, MD, David Geffen School of Medicine at UCLA; Bruce Psaty, MD, PhD, University of Washington) and 1 federal partner (the National Institutes of Health). Comments from reviewers were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

Editorial Disclaimer: This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to JAMA.

REFERENCES


42. Messenger B, Li D, Nasiri K, Carr JJ, Blankstein R, Budoff MJ. Coronary calcium scans and radiation exposure in the multi-ethnic study of


93. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007;115(7):928-935. doi:10.1161/CIRCULATIONAHA.106.672402
