Risk Prediction in Cardiovascular Medicine

Assessing the Role of Circulating, Genetic, and Imaging Biomarkers in Cardiovascular Risk Prediction

Thomas J. Wang, MD

Never make forecasts. Especially about the future.

—Samuel Goldwyn

The use of biomarkers to augment traditional cardiovascular risk prediction has attracted considerable attention in the past decade. This interest has been fueled by the realization that traditional risk factors do not identify everyone who will eventually develop cardiovascular disease. This has been accompanied by the emergence of potential screening tests such as high-sensitivity C-reactive protein (CRP), single-nucleotide polymorphism arrays, and coronary calcium scanning.

Nonetheless, whether these biomarkers are ready for routine clinical use, to be measured alongside existing tests such as cholesterol, is controversial. Recently, the US Preventive Services Task Force concluded that “current evidence is insufficient to assess the balance of benefits and harms of using the nontraditional risk factors.” This conclusion underscores gaps in the current literature, as well as the inherent challenge of predicting the occurrence of future cardiovascular events in generally healthy adults. However, the use of newer screening tests for cardiovascular risk assessment has been endorsed in some quarters. For instance, the state of Texas passed legislation in 2009 mandating the coverage of screening coronary calcium scans in older adults. Furthermore, the Food and Drug Administration recently expanded the labeling of a statin drug, rosuvastatin, to include older individuals with elevated CRP levels in addition to conventional risk factors.

The present article reviews the case for and against the use of newer biomarkers for cardiovascular risk prediction. There are literally hundreds of circulating, genetic, and imaging biomarkers that have been proposed or evaluated in the cardiovascular literature. Several comprehensive reviews exist on the topic. Rather than attempting to provide a systematic overview, this article focuses on a few of the most illustrative biomarkers. Particular emphasis will be given to placing these biomarkers in a pathophysiological context (what stage of disease is represented?), addressing the utility of a multimarker approach (does it help to combine biomarkers?), and proposing priorities for future research in this area (where do we go from here?).

Why Do We Need Biomarkers?

Cardiovascular disease remains the leading cause of death in the United States, a fact that underscores the importance of primary prevention. The success of preventive measures depends in part on the accurate identification of individuals who are at risk for future cardiovascular events (risk prediction). Traditionally, risk prediction has relied on assessment of risk factors such as hypertension, diabetes mellitus, hyperlipidemia, and smoking. Indeed, the approach to risk prediction has changed relatively little since Dr William Kannel coined the phrase factors of risk in 1961 to describe these predisposing conditions.

Despite the value of traditional risk factors, as many as half of individuals who develop coronary heart disease have only 1 or none of these risk factors. Indeed, individuals with few clinical risk factors are the least likely to be targeted for preventive therapies, but as a group they experience the largest number of cardiovascular events. Thus, there is an important need for new ways to improve on the information obtained from traditional risk assessment so that preventive measures can be applied to those who are the most likely to experience events.

Efforts to address the problem of identification have largely focused on the use of new measurements, or biomarkers, that could be added to traditional ones such as blood pressure and cholesterol. In 2001, a National Institutes of Health working group proposed a definition of a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.” This definition highlights the fact that biomarkers are not limited to blood tests, although that is the context in which the term is most frequently used in the cardiology literature. Thus, a biomarker could include any representation of a biological process, including circulating molecules, genetic markers, cellular markers, results of imaging, or findings on physical examination.

Furthermore, which of these biomarkers is most informative is partly a function of the phase in the disease process (Figure 1). It is now recognized that overt cardiovascular disease is typically preceded by a long period of subclinical cardiovascular disease. For instance, subclinical atheroscle-
Discrimination
At least 3 criteria other than statistical significance have been proposed in the evaluation of new biomarkers: discrimination, calibration, and reclassification. Discrimination refers to the ability of a test to distinguish those who will get disease from those who will not. It is the basis for the most widely used metric for new screening tests: the C statistic or area under the receiver operating characteristic curve (AUC). The AUC incorporates 2 measures of the accuracy of a screening or diagnostic test: its sensitivity (ability to detect disease when it is present [eg, true positive rate]) and its specificity (ability to exclude disease when it is absent [eg, true negative rate]). Values for the AUC or C statistic range from 0.5 (uninformative model) to 1.0 (perfect discrimination). In Framingham data, a model for coronary heart disease with age and sex alone has a C statistic of ~0.68. The Framingham Risk Score, which is based on traditional cardiovascular risk factors, yields a C statistic of ~0.75. Thus, one approach to evaluating new cardiovascular biomarkers is to assess whether they raise the C statistic above what is obtainable with the Framingham Risk Score alone.

Although sensitivity and specificity are traditionally regarded as “intrinsic” attributes of a test, it is well known that estimates of sensitivity and specificity can vary according to the characteristics of the population. This phenomenon, known as “spectrum bias” or “spectrum effect,” can influence estimates of discrimination. For biomarker studies, spectrum bias is often manifested by divergent estimates of the AUC in high- versus low-risk individuals or in older versus younger individuals.

Reclassification
A drawback of global measures of calibration such as the Hosmer-Lemeshow statistic is that they give equal weight to all combinations of predicted and observed risk. Most of the difficult treatment decisions apply to individuals in a narrow range of risk (for instance, those who are “intermediate risk”). Recently, an alternate approach has been proposed to capture the impact of new biomarker tests on clinical decision making, namely, to see whether the addition of a new risk marker results in a substantial proportion of individuals being...
The most appropriate setting in which to use reclassification to evaluate risk models is for conditions in which risk strata are clearly defined and closely linked to treatment decisions. For coronary heart disease, risk strata are based on recommendations of the Third Adult Treatment Panel of the National Cholesterol Education Program. Low-, intermediate-, and high-risk categories are based on having predicted 10-year risks of coronary heart disease <10%, 10% to <20%, or ≥20%, respectively. It is recommended that individuals in the high-risk category be considered for the same therapies as those with known coronary heart disease, including the use of statins to reduce low-density lipoprotein (LDL) cholesterol to <70 mg/dL.

The US Preventive Services Task Force has endorsed reclassification as an important indicator of the clinical value of that test. The US Preventive Services Task Force recommendations focus specifically on intermediate-risk individuals on the basis of the premise that high-risk individuals should automatically be targeted for preventive measures (including pharmacological therapy), whereas low-risk individuals do not need specific intervention. Reclassifying intermediate-risk individuals up (to high risk) or down (to low risk) could lead to a clear treatment decision. A similar approach has been advocated by previous consensus conferences.

Table 1 summarizes the strengths and limitations of different criteria for evaluating new biomarkers. As shown, although reclassification is the most clinically intuitive of the criteria, it has limitations as well. Clinical reclassification is most important when the categories correspond to established clinical criteria. Although a predicted risk of 20% in 10 years is the conventional threshold for justifying the use of pharmacological therapies to achieve very low LDL targets (<70 mg/dL), this threshold is consensus based rather than evidence based. Indeed, one could argue on the basis of recent primary prevention trials that lower thresholds for treatment may be warranted. Even when there is agreement about thresholds, not all movements across thresholds lead to clinically meaningful changes. For instance, movement from low risk to intermediate risk, or vice versa, likely results in no change in the treatment decision and arguably should not be included in the assessment of proportion reclassified.

Furthermore, simply quantifying the proportion of individuals reclassified across a given threshold ignores whether that reclassification was correct. One could reclassify 100% of intermediate-risk individuals by using a coin flip to determine who is reclassified up versus down, but such a strategy would yield no net benefit (or could be detrimental). Accordingly, several authors have emphasized the importance of characterizing the accuracy of reclassification either by comparing observed rates to predicted strata or using more formal measures. As an example of the latter, Pencina and

Table 1. Assessing the Incremental Predictive Value of New Biomarkers, Selected Criteria

<table>
<thead>
<tr>
<th>Metric(s)</th>
<th>Attribute</th>
<th>Description</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR, HR</td>
<td>Association</td>
<td>Basic measure of strength of statistical association between marker and outcome</td>
<td>Easily understood Significant OR or HR a necessary criterion for a new risk marker</td>
<td>High OR or HR not adequate to ensure predictive accuracy Does not take into account prevalence of positive marker values</td>
</tr>
<tr>
<td>Likelihood ratio test, Bayes information criterion</td>
<td>Model fit</td>
<td>Assesses improvement in model fit from addition of new variable(s)</td>
<td>Sensitive to information from new predictors</td>
<td>Positive result not adequate to ensure predictive accuracy</td>
</tr>
<tr>
<td>C statistic, AUC</td>
<td>Discrimination</td>
<td>Captures the sensitivity and specificity of a marker across the full range of cut points</td>
<td>Easily understood No reliance on specific thresholds or categories</td>
<td>Difficult for new predictors to raise the C statistic when existing variables discriminate well</td>
</tr>
<tr>
<td>Hosmer-Lemeshow statistic</td>
<td>Calibration</td>
<td>Assesses concordance between predicted and observed event rates in longitudinal data</td>
<td>Assesses accuracy of predicted risks, which are the basis for clinical decisions</td>
<td>Relatively insensitive to differences between models Provides a global measure, even though clinical decisions may be based on predicted risks within a relatively narrow range</td>
</tr>
<tr>
<td>Proportion reclassified</td>
<td>Reclassification</td>
<td>Assesses the proportion of individuals moved between risk categories by addition of new biomarker(s)</td>
<td>Clinically relevant when risk categories are linked to treatment decisions</td>
<td>Sensitive to changes in the number of risk categories and choice of cut points Does not assess accuracy of reclassification Includes reclassifications that are unlikely to affect clinical decisions</td>
</tr>
<tr>
<td>NRI</td>
<td>Reclassification</td>
<td>Assesses the proportion of individuals reclassified correctly by the addition of new biomarker(s)</td>
<td>Clinically relevant when risk categories are linked to treatment decisions Incorporates information on the accuracy of reclassification</td>
<td>Sensitive to changes in the number of risk categories and choice of cut points Gives some weight to reclassifications that are unlikely to affect clinical decisions</td>
</tr>
</tbody>
</table>

OR indicates odds ratio; HR, hazard ratio. For references, see text and prior reviews on this topic. The list contains the metrics used most commonly in the biomarker literature. Other metrics that have been described for risk models include the integrated discrimination improvement and the reclassification calibration statistic.
colleagues\textsuperscript{25} have proposed a statistic termed the net reclassification improvement (NRI). The NRI is calculated by splitting the sample into those who develop events and those who do not. In the event group, those who are up-classified by a new biomarker are considered correctly reclassified, whereas those who are down-classified are considered incorrectly reclassified. A similar tally is made for the nonevent group. The net number of correct reclassifications (correct minus incorrect, in both groups) yields the NRI. A coin flip would have a NRI of 0\% (on average) because it would lead to an equal number of correct and incorrect reclassifications.

The subsequent sections examine the predictive utility of circulating, imaging, and genetic biomarkers for cardiovascular disease. Because the value of biomarker screening is ultimately determined by whether addition of these tests alters patient outcomes, the article concludes with discussion of biomarker-guided therapies and cost-effectiveness.

**Are Circulating Biomarkers Useful for Risk Prediction?**

Numerous circulating biomarkers have been related to cardiovascular risk.\textsuperscript{3,4} These biomarkers represent a variety of pathophysiological pathways that promote atherogenesis or cardiac dysfunction, including inflammation (such as CRP, interleukin-6, lipoprotein-associated phospholipase A2 [LpPLA\textsubscript{2}]), oxidative stress (oxidized LDL, nitrotyrosine), lipid metabolism [lipoprotein(a)], thrombosis (plasminogen activator inhibitor-1, D-dimer), endothelial dysfunction (homocysteine, urinary microalbuminuria), hemodynamic stress (natriuretic peptides), and cardiomyocyte injury (cardiac troponins).\textsuperscript{37} The best evidence for the majority of these biomarkers is their association with future events. Data addressing how these biomarkers perform in terms of discrimination, calibration, or reclassification are far more limited.

**The Evidence for CRP**

The biomarker for which the largest body of data exists is CRP, an acute-phase reactant of hepatic origin. In 1997, Ridker and colleagues\textsuperscript{38} reported the results of a nested case-control study showing that baseline CRP concentrations were strongly associated with future myocardial infarction. Since then, multiple epidemiological reports have been published, all generally supporting an association between CRP concentrations and future cardiovascular risk, although estimates of effect vary.\textsuperscript{38} A meta-analysis of 22 studies was performed recently by the US Preventive Services Task Force and showed that CRP concentrations of $\geq 3.0$ mg/dL were associated with roughly 60\% excess risk of incident coronary heart disease (risk ratio, 1.60; 95\% confidence interval, 1.43 to 1.78).\textsuperscript{39} Findings were similar when the meta-analysis was restricted to the 10 cohort studies judged to be “good quality” by predefined criteria.

Despite the robust statistical association between CRP and cardiovascular events, evidence from multiple studies indicates that CRP measurements provide only modest improvements in predictive accuracy.\textsuperscript{40–42} The addition of CRP raises the C statistic by 0.0 to 0.02 compared with standard risk models, signifying little change in discrimination. Similarly, studies have found either modest\textsuperscript{43} or absent\textsuperscript{41,42} improvements in model calibration with the addition of CRP.

The value of CRP with regard to reclassification is uncertain. In the Women’s Health Study, addition of CRP to standard risk models reclassified 20\% of intermediate-risk individuals (baseline 10-year predicted risk, 6\% to 20\%).\textsuperscript{43} Nearly three quarters of these individuals were reclassified downward to low risk. Thus, only 4\% of the intermediate-risk group (<0.5\% of the overall sample) were reclassified from intermediate risk to high risk, the transition most likely to influence management in individuals not already on treatment. Similar results were noted in the Swedish Malmo Diet and Cancer Study with the use of a biomarker score that included CRP.\textsuperscript{42} Overall, 16\% of intermediate-risk participants were reclassified, with three quarters (13\% of the intermediate-risk group) reclassified downward. Thus, the proportion of intermediate-risk individuals reclassified upward was only 3\%.

The accuracy of reclassification with CRP has been assessed in several studies with the use of the NRI. In the Women’s Health Study, the NRI was of modest magnitude (5.7\%).\textsuperscript{43} In the Framingham Offspring Study, the NRI was 10.9\% for coronary heart disease and 5.6\% for overall cardiovascular disease.\textsuperscript{44} On the other hand, the investigators from the Malmo Diet and Cancer Study found the NRI for CRP to be nonsignificant for both end points (NRI <2\%).\textsuperscript{42}

Although it is widely accepted that vascular inflammation plays a role in atherogenesis, there is little evidence that CRP specifically contributes to this process.\textsuperscript{45} Transgenic animals with overexpression of CRP do not exhibit increased atherosclerosis.\textsuperscript{46–50} Furthermore, recent population genetic studies utilizing the “mendelian randomization” concept have failed to support a direct effect of CRP on cardiovascular risk.\textsuperscript{51,52} Nonetheless, it is important to emphasize that causality is not a requirement for measuring a biomarker, provided that the biomarker adds predictive information that aids clinical management.

**Biomarkers Not Linked to Inflammation: The Example of B-Type Natriuretic Peptide**

Whereas elevations in CRP may reflect pathogenic processes (inflammation) that give rise to cardiovascular disease, other biomarkers may arise as a consequence of cardiovascular injury or stress.\textsuperscript{53} One example is B-type natriuretic peptide (BNP), which has been recognized as a sensitive marker of acute heart failure but has been appreciated only recently as a predictor of future cardiovascular disease. BNP is a member of a family of peptide hormones, the natriuretic peptides, which have vasodilatory, natriuretic, and antihypertrophic properties. BNP is synthesized predominantly by cardiac tissue. Stretch of cardiac myocytes leads to rapid activation of the BNP gene,\textsuperscript{54} and translation of the gene transcript for BNP produces pre-proBNP, which is then cleaved to form the 108–amino acid peptide precursor proBNP. ProBNP is cleaved into the biologically inactive N-terminal-proBNP (NT-proBNP) and the biologically active c-terminal fragment (mature BNP).
Circulating NT-proBNP and BNP are particularly elevated in individuals with established cardiovascular disease. Concentrations in apparently healthy individuals are much lower, but it is now known that variation in BNP within the “normal range” can still have significant prognostic value. In an investigation from the Framingham Offspring Study, BNP levels >20 pg/mL were associated with 60% to 200% increased risks of future cardiovascular events, stroke, heart failure, and all-cause mortality. Similar findings have been observed in other prospective cohorts. A recent meta-analysis of 40 cohort studies that measured BNP or NT-proBNP found a summary relative risk of 2.82 (95% confidence interval, 2.40 to 3.33) for incident cardiovascular disease in those with BNP levels in the top versus bottom tertile. Interestingly, BNP (or NT-proBNP) concentrations predict future coronary disease nearly as well as they predict future heart failure. It is unclear whether this represents low-grade increases in cardiac wall stress secondary to subclinical ischemia or a direct effect of ischemia or vascular disease on natriuretic peptide release.

The predictive accuracy of BNP or NT-proBNP has been examined in only a few studies, some of which are restricted to high-risk subjects. In the Heart Outcomes Prevention Evaluation (HOPE) study, which included predominantly individuals with prior vascular disease, addition of NT-proBNP to risk models improved discrimination, raising the C statistic from 0.65 (with traditional risk factors) to 0.69 (with traditional risk factors plus NT-proBNP). In the Heart and Soul study, a secondary prevention cohort, NT-proBNP raised the C statistic from 0.73 to 0.80. However, in 3 primary prevention cohorts, the addition of NT-proBNP to cardiovascular risk models led to smaller changes in the C statistic (ranging from 0.01 to 0.03). In 2 of these studies, reclassification was assessed with the NRI and found not to be significant for cardiovascular events or cardiovascular mortality. Spectrum bias could be 1 explanation for the discrepant findings with NT-proBNP between high-risk and low-risk populations.

Interestingly, in studies that have compared BNP (or NT-proBNP) with CRP “head to head,” BNP has demonstrated consistently stronger associations with vascular events. For instance, in the HOPE study, NT-proBNP but not CRP was retained in stepwise selection models testing multiple biomarkers, and only NT-proBNP raised the C statistic significantly. A similar pattern has been observed in primary prevention studies, in which BNP (or NT-proBNP) has outperformed CRP in the prediction of cardiovascular events, coronary events, and cardiovascular mortality.

Other Biomarkers
As noted above, a wide range of other circulating biomarkers has been studied for their ability to predict risk of future cardiovascular events. One of the biomarkers that has generated recent interest is LpPLA₂. This enzyme is expressed by inflammatory cells and is carried in the circulation bound to LDL. The action of LpPLA₂ on membrane phospholipids leads to the production of proinflammatory factors. A recent meta-analysis of 32 prospective studies showed a linear relation between LpPLA₂ activity and mass and vascular risk in individuals with and without baseline cardiovascular disease. However, there are few data with regard to the ability of LpPLA₂ to improve risk discrimination, calibration, or classification.

Another inflammatory biomarker that has shown robust associations with cardiovascular risk is interleukin-6. Interleukin-6 is a proinflammatory cytokine that is responsible for activating downstream molecules including CRP. Notably, the association of interleukin-6 with future events is not significantly attenuated by adjustment for CRP or fibrinogen. As with LpPLA₂, there are few data on the ability of interleukin-6 to improve risk prediction with the use of current metrics.

Interest in the novel lipid biomarker lipoprotein(a) has also increased recently, in part due to mendelian randomization data suggesting a causal role for this lipoprotein in coronary heart disease. Although the majority of observational studies that have assessed the relation of lipoprotein(a) with cardiovascular events have found a significant association, it is unknown whether measurement of this biomarker improves discrimination, calibration, or classification.

In summary, existing circulating biomarkers lead to relatively modest enhancements in risk prediction. Where present, the evidence for benefit is restricted largely to their ability to reclassify a proportion of individuals to a higher or lower risk category rather than any demonstrable effect on discrimination or calibration. Whether these reclassifications are valuable for guiding treatment is discussed in greater detail below.

Are Genetic Biomarkers Useful for Risk Prediction?
The sequencing of the human genome has fueled an explosion of knowledge about the genetic underpinnings of complex diseases such as cardiovascular disease. An active area of discovery has been the identification of common genetic variants (eg, present in more than a few individuals per 100) that have been reproducibly related to myocardial infarction in large genome-wide association studies. More than 10 validated genetic variants have been reported to date, and the number continues to grow.

Genetic risk markers are present at birth and thus can be ascertained well before an individual has developed conventional risk factors for cardiovascular disease (primordial prevention). This possibility brings a new dimension to the potential applications of risk prediction in large populations. Nonetheless, the considerations that are important for assessing circulating biomarkers also apply to genetic markers. Despite the strong, statistically significant associations between genetic variants and myocardial infarction risk, effect sizes are typically small, ranging from 10% to 30% increased risk per copy of the risk allele. These effects would be expected to translate into modest improvements, if any, in risk discrimination or calibration. This impression is supported by an analysis from the Women’s Health Study cohort examining single-nucleotide polymorphisms at chromosome 9p21 and the risk of cardiovascular disease in 22,129 women. Common variants at 9p21 exhibit the strongest and most repro-
ducible associations with cardiovascular disease in genome-wide association studies. However, the addition of 9p21 variants to a conventional risk factor model in the Women’s Health Study did not change the C statistic or improve calibration. Only 2.7% of women were reclassified, and the NRI was modest (2.7%).

There are ongoing efforts to achieve better performance by combining information from multiple genetic markers (see section on Multiple Biomarkers). Nonetheless, the strongest common variant signals have probably already been identified by existing genome-wide association studies. Relative risks per allele for the most recently discovered variants are only \( \approx 1.1 \), in contrast to 1.25 to 1.30 for variants at chromosome 9p21. Accordingly, although the number of genetic markers is increasing, the incremental gain from each of the markers is decreasing. Alternate approaches exist, such as testing for less common or rare genetic variants that may harbor stronger associations. The relevance of such approaches for widespread screening is uncertain at best, however.

**Are Imaging Biomarkers Useful for Risk Prediction?**

Radiological advances have opened the possibility that noninvasive vascular imaging could be performed in broad populations to identify asymptomatic individuals at particularly high risk for cardiovascular events. In contrast to some circulating biomarkers and all genetic biomarkers, these imaging tests detect existing disease, typically signs of atherosclerosis (Figure 1). This might have the advantage of improved predictive accuracy, particularly in the short to medium term, because individuals with subclinical disease may be closer to having overt clinical events compared with those who simply have a predisposition for disease. On the other hand, it might be preferable to implement preventive measures before the development of even subclinical disease. Furthermore, some atherosclerotic plaques are below the detection limits of current imaging modalities, and the likelihood of plaque rupture is not necessarily correlated with the size of the plaque or likelihood of being detected by imaging tests. The most widely studied imaging tests of subclinical atherosclerosis are coronary calcium scanning and carotid ultrasound. Other modalities such as magnetic resonance imaging may have future promise but are currently impractical for consideration of widespread testing.

**Coronary Calcium Scores**

Coronary calcium may be detected by electron beam computed tomography or multidetector computed tomography. Multidetector computed tomography is a newer technology with high resolution and slightly more radiation exposure, whereas electron beam computed tomography has the largest body of existing evidence. Both may be used to compute coronary calcium scores (typically, with the use of the method of Agatston et al\(^{[2,72]}\)). Computed tomographic angiography is another noninvasive test based on this technology but is of less direct relevance to screening given the use of intravenous contrast dye.

A number of observational studies have documented strong associations between coronary calcium scores and the risk of incident cardiovascular events. Several comprehensive reviews of this literature have been published. Although many of the earlier studies were based on individuals referred for electron beam computed tomography or multidetector computed tomography, and thus susceptible to referral bias, later studies examined largely unsampled populations. Most reports indicate a relative risk of coronary events of \( \geq 2 \) for those with any coronary calcium compared with those with no coronary calcium and a relative risk of \( > 4 \) for those with coronary calcium scores \( > 100 \).

The available evidence also suggests that coronary calcium scores may improve risk discrimination. In 6698 individuals enrolled in the Multi-Ethnic Study of Atherosclerosis (MESA), coronary calcium scores were strong predictors of cardiovascular and coronary events and improved the C statistic from 0.77 (with traditional risk factors) to 0.81 (traditional risk factors plus coronary calcium score). Similarly, in 2028 asymptomatic individuals in the Rotterdam Study, coronary calcium scores raised the C statistic for coronary events from 0.72 to 0.76.

There is also increasing evidence that coronary calcium scores improve risk classification, particularly for intermediate-risk individuals. In MESA, calcium scores reclassified 54% of intermediate-risk individuals, including 16% who were upclassified to high risk. The overall NRI was 25%, with an approximately equal proportion of correct up-reclassifications and correct down-reclassifications in the intermediate-risk group. In the Rotterdam Study, 52% of intermediate-risk individuals were reclassified, 22% as high risk, and the overall NRI was 14%. Similar findings were obtained from another large, population-based cohort, the Heinz Nixdorf Recall Study, with >60% of intermediate-risk individuals reclassified, 27% as high risk. The NRI in that study was 22%.

In contrast to other screening biomarkers under consideration, there is the potential for harm from coronary calcium scanning related to the use of ionizing radiation. With current technologies, the magnitude of radiation exposure is relatively small and decreasing. The radiation from a single coronary calcium scan is approximately equal to the environmental radiation that the average person experiences in 1 year. Nonetheless, because the adverse effects of radiation exposure are cumulative, expanded use of coronary calcium scanning for screening large populations could have public health consequences that would have to be weighed against the benefits of cardiovascular prevention.

**Carotid Intima-Media Thickness**

Another established technology for assessing subclinical atherosclerosis is carotid ultrasound. Clinical studies have focused in particular on carotid intima-media thickness (IMT), the combined thickness of the intimal and medial layers of the carotid artery, which is typically measured at the far wall of the common or internal carotid arteries. Because elevated IMT probably predates atherosclerotic plaque, it can be viewed as an earlier biomarker of subclinical atherosclerosis than coronary calcium.
Multiple epidemiological studies have examined the association of carotid IMT with cardiovascular risk (reviewed in Reference 86). Although estimates of effect vary widely, most studies with adequate adjustment for traditional risk factors suggest an independent association between carotid IMT and cardiovascular events. A meta-analysis of 8 large, prospective studies found that each 0.1-mm increment in carotid IMT was associated with a 10% to 15% increase in the risk of myocardial infarction.87 However, there are few data showing that carotid IMT improves measures of predictive performance. Folsom and colleagues83 found that carotid IMT led to a minimal increase in the C statistic for predicting cardiovascular or coronary events in the MESA cohort. Whether carotid IMT improves calibration or reclassification is not known.

Presently, the use of carotid IMT is restricted to research settings. It is one of the most commonly ascertained noninvasive measures in cardiovascular epidemiology studies. It has also gained popularity as a surrogate end point in randomized trials of new cardiovascular therapies.88,89 Nonetheless, even if clinical studies eventually support the use of IMT for screening selected populations, it is important to recognize that most research studies utilize experienced testing sites and centralized reading centers. It is unclear whether techniques for measuring carotid IMT could be sufficiently standardized between laboratories to allow for widespread clinical screening, particularly when one considers that submillimeter differences in IMT separate low-risk and high-risk groups.

**Tests of Vascular Function**

Alterations in vascular function may precede anatomic evidence of atherosclerosis, suggesting noninvasive measures of vascular function as potential early biomarkers of disease. One newer tool, vascular tonometry, assesses pathophysiological changes in arterial walls by providing measures of arterial wall compliance at different sites. Reduced compliance in large arteries reflects structural remodeling in response to elevated pressure, aging, or disease. In smaller arteries, changes in compliance reflect altered vascular smooth muscle function. Aortic pulse wave velocity is the best-studied measure of large-artery stiffness. A recent meta-analysis of 18 studies showed that high pulse wave velocity was associated with ≈2-fold risks of cardiovascular events, cardiovascular mortality, and all-cause mortality, although there was a tendency toward smaller relative risk estimates in cohorts at low baseline risk. Data from a few studies suggest that pulse wave velocity may modestly improve discrimination and reclassification.

Another potential biomarker is flow-mediated dilation assessed with the use of ultrasound of the brachial artery. Because vasodilation after shear stress is a nitric oxide–dependent process, flow-mediated dilation provides an index of endothelial function. Reduced flow-mediated dilation is present in individuals with hypertension, hyperlipidemia, and other cardiovascular risk factors. A number of studies have shown that reduced flow-mediated dilation is associated with cardiovascular risk in individuals with varying baseline risk.93–100 Nonetheless, its predictive value in relatively low-risk individuals remains to be established. In a recent report from MESA, flow-mediated dilation did not improve the C statistic and inappropriately reclassified 23% of individuals who experienced events.98 Furthermore, brachial ultrasound requires significant expertise and is currently only performed in research settings, making it impractical for use as a screening tool.

Although not technologically an imaging test, the ankle-brachial index is a simple, noninvasive measure that provides information on the presence of peripheral arterial disease. An abnormal ankle-brachial index could serve as a biomarker of overall cardiovascular risk because individuals with peripheral arterial disease typically have atherosclerosis in other arterial beds as well. Indeed, meta-analysis of 16 cohort studies showed that an ankle-brachial index ≤0.90 was associated with an approximate doubling of the risk of cardiovascular mortality and major coronary events, even after adjustment for Framingham Risk Score.101 This meta-analysis suggested that 4% of intermediate-risk men and 10% of intermediate-risk women could be reclassified by adding ankle-brachial index to the Framingham Risk Score. The majority of reclassification in men was from a higher to a lower category, whereas the opposite was true in women, perhaps attributable in part to the far higher proportion of women in lower-risk categories in these studies (another example of spectrum bias).

In summary, there are accumulating data on a variety of imaging biomarkers of cardiovascular disease. Thus far, the strongest associations with events have been observed for coronary calcium scores, perhaps because coronary calcification reflects more established atherosclerosis, derives from the coronary vasculature, and can be measured reproducibly. The available data also suggest that coronary calcium scores improve risk discrimination and reclassify a proportion of intermediate-risk individuals. Nonetheless, the clinical implications of elevated calcium scores and the risks associated with widespread use of coronary calcium scanning have not been fully elucidated.

**What Is the Utility of Combining Multiple Biomarkers?**

The relatively modest performance of individual biomarkers for risk prediction raises the question of whether multiple biomarkers could be combined to improve performance. The “multimarker approach” has been tested in several studies, primarily with the use of circulating biomarkers. There are fewer data incorporating imaging into multimarker algorithms and also few data on the use of circulating, genetic, and/or imaging biomarkers in combination.

**Existing Evidence**

An investigation from the Framingham Heart Study examined 10 biomarkers (9 circulating and 1 urinary) for cardiovascular risk prediction in ≈3000 individuals.13 The 2 biomarkers most strongly associated with cardiovascular events were BNP and the ratio of urinary albumin to creatinine. The addition of the multiple biomarkers, either BNP and ratio of
The hypothesis that combining multiple genetic markers can improve risk prediction was tested by investigators from the Malmo Diet and Cancer study. A score comprising 9 lipid-associated single-nucleotide polymorphisms was incorporated into risk models and proved to be a significant predictor of cardiovascular events even after adjustment for measured lipid levels. This score did not improve the C statistic but was associated with a moderate improvement in classification accuracy (NRI, 6.2%).

**Future Prospects: How Many Markers Would It Take?**

Thus, studies that have examined the use of up to 10 biomarkers simultaneously suggest, at best, modest improvements in risk prediction. This observation is a testament to the value of traditional risk factors as much as it is an indication of the difficulty of predicting future events that are multifactorial in origin and inherently stochastic. Consequently, is it realistic to expect that any set of biomarkers (circulating, imaging, or genetic) can substantially improve risk prediction above and beyond traditional risk scores?

One way to consider this question is shown in Figure 2. The figure depicts the results of a simulation based on adding 1 to 100 hypothetical biomarkers to a traditional risk model (written communication, M. Pencina, PhD, Boston University, October 14, 2010). Each of the hypothetical biomarkers has a similar magnitude of association with cardiovascular events (the outcome) as a marker such as CRP or BNP. The increment in C statistic is shown, from a baseline of 0.75 (the approximate value with traditional risk factors alone) to the value with risk factors and biomarkers combined. A key determinant of improvement in the C statistic is the degree of correlation between biomarkers. With a set of biomarkers that has an average marker-marker correlation of r=0.4 (moderately correlated), >50 biomarkers are needed before the C statistic is increased by 0.05. In contrast, when the average marker-marker correlation is r=0.05 (weakly correlated), <10 biomarkers are needed to raise the C statistic by 0.05.

Thus, it is theoretically possible to improve the performance of risk models with a relatively small number of biomarkers, provided that the biomarkers are weakly correlated or uncorrelated with each other. This concept is intuitive but suggests a dilemma. Studies tend to focus on biomarkers derived from the pathways implicated in atherogenesis, such as inflammation or thrombosis. This pathway-based approach is valuable for elucidating biology, but, paradoxically, it may be limited when it comes to identifying clinically useful biomarkers. When a biomarker of a known pathway is already being measured, more biomarkers from that pathway are not likely to yield substantial gains. For instance, because CRP is easy to measure and widely available, there is comparatively little value to characterizing a large number of other cytokines or inflammatory markers, at least for the purpose of risk prediction.

Where does one find “uncorrelated” biomarkers, outside of already characterized pathways? The opposite of a pathway-based approach is an “unbiased” approach, in which global panels of biomarkers are compared in individuals with and
without disease. Such analyses are now possible with newer technologies that are capable of systematically profiling hundreds or thousands of molecules in a biological specimen. The application of such technologies (proteomics, metabolomics, expression profiling) for biomarker discovery has been reviewed in more detail elsewhere. Studies in this area are still hampered by problems such as multiple testing, confounding, and overfitting. As a result, there can be a fine line between an “unbiased” biomarker study and a “fishing expedition.” That said, similar issues affected early studies in population genetics. As large-scale, collaborative approaches extend from genetics to other types of biomarker studies, the possibility of robust and clinically useful biomarker discoveries is enhanced.

Clinical Implications

Effectiveness of Biomarker-Guided Therapies

Ultimately, the most important question for assessing new biomarkers is whether they lead to changes in patient management and improvements in outcome. The gold standard for establishing the effectiveness of biomarker-guided strategies is randomized trials. However, there are few examples of such trials in cardiology, particularly in the primary prevention setting. There are several possibilities for the design of biomarker trials, as depicted in Figure 3A and 3B and as reviewed previously. In 1 design, all subjects have measurement of a biomarker, and both biomarker subgroups (those with a “positive” result and those with a “negative” result) undergo randomization to treatment A or treatment B (or placebo) (Figure 3A). If treatment A is an effective therapy, this design permits a test of whether the biomarker stratification identifies those who are more likely to benefit from the therapy (eg, biomarker-by-treatment interaction). The negative biomarker arm is necessary because it is possible that treatment A is effective (eg, better than treatment B) in both biomarker groups, which can only be determined by including both groups. This scenario can occur either in the absence of a biomarker-by-treatment interaction (eg, treatment A is equally effective in both positive and negative biomarker groups) or in the presence of a biomarker-by-treatment interaction (eg, treatment A is more effective in the positive biomarker group but still better than treatment B in everyone). Thus, even in the presence of a biomarker-by-treatment interaction, measurement of the biomarker may not be necessary to guide therapy. Although the comparison shown in Figure 3A can be performed post hoc, by randomizing all subjects and measuring the biomarker later, implementing it prospectively ensures balanced treatment within marker groups and reduces the risk of false-positives inherent in subgroup analyses.

The study design outlined above is an indirect test of biomarker-guided strategies because it does not actually provide a measure of how often patient management would change or patient outcomes would improve. In addition, because the biomarker test is performed in everyone, knowledge of the biomarker result could affect other management decisions or other outcomes of interest. A more direct assessment is shown in Figure 3B and incorporates an earlier step of randomization to biomarker measurement versus no biomarker measurement. This design is frequently used in oncology trials.

Example: The JUPITER Trial

In cardiology, there has been only 1 large-scale randomized trial involving novel biomarker-directed therapy in primary prevention: Justification for Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER). JUPITER enrolled middle-aged to elderly individuals treated with rosuvastatin who had baseline LDL <130 mg/dL and CRP >2 mg/dL. The incidence of cardiovascular events...
was reduced 44% by treatment with rosuvastatin compared with placebo.

Although JUPITER was an important trial, it is apparent that it fits neither of the study designs depicted in Figure 3A and 3B. Because there was no randomization to biomarker measurement, JUPITER is closest to the “indirect” approach shown in Figure 3A. However, a key difference is that there was no negative or low biomarker subgroup in JUPITER, making it impossible to exclude the possibility that rosuvastatin therapy is beneficial for all older individuals, regardless of baseline CRP. Accordingly, JUPITER cannot be considered either a “biomarker trial” or “CRP trial” because the utility of CRP screening was not specifically evaluated.

On the basis of the JUPITER results, can we conclude that CRP screening is either necessary or sufficient to determine which individuals will benefit from statin therapy in primary prevention? Two observations suggest that an elevated CRP may not be a necessary criterion. First, on the basis of traditional risk factors alone, the average participant in JUPITER was at intermediate risk, with an average Framingham Risk Score of 11.6%. The mean LDL level was ∼100 mg/dL. Meta-analyses of data from prior statin trials by the Cholesterol Treatment Trialists suggest that the benefit of statins extends to those with baseline LDLs as low as 80 mg/dL. Thus, the results of JUPITER may have been predictable on the basis on prior trials that did not include CRP testing as an entry criterion. Second, in 1 of the prespecified subgroup analyses in JUPITER, the efficacy of rosuvastatin was examined separately in those above and below the median CRP level (4.2 mg/L). Surprisingly, those with higher CRP levels (>4.2 mg/L) had a lower relative risk reduction from rosuvastatin therapy than those with lower CRP levels (relative risk reduction 29% versus 58% in those with higher versus lower CRP levels, respectively; P for interaction = 0.015). In epidemiological studies, the relationship between CRP and cardiovascular risk is linear, with no evidence of a threshold at levels observed in the general population. Thus, even though JUPITER excluded those with below-average CRP levels, the absence of a positive treatment-by-biomarker interaction over the range of levels studied suggests that elevated CRP was not an important determinant of relative benefit from rosuvastatin.

There is also no evidence that an elevated CRP is sufficient to identify candidates for statin therapy. In JUPITER, there was no evidence of benefit from rosuvastatin in individuals who met the age and CRP criteria but had no other traditional cardiovascular risk factors. For these participants, the relative risk with rosuvastatin was 0.91 (95% confidence interval, 0.56 to 1.46). Although post hoc analyses should be interpreted cautiously, the results are consistent with the large body of evidence that underscores the importance of global risk stratification with assessment of traditional risk factors.

More General Implications From JUPITER

Do the JUPITER results have any implications for screening with cardiovascular biomarkers in general? The relative risk reduction with rosuvastatin was remarkably consistent across subgroups, except for those at very low risk as noted above. Thus, for most statin candidates, the total number of events prevented is primarily a function of the baseline absolute risk. Accordingly, the number of events prevented in low-risk individuals may not be high enough to warrant the costs or toxicity of therapy, including the potential risk of diabetes mellitus, one of the adverse safety signals in JUPITER and other statin trials. Even if we assume that true relative risk reductions were constant in all risk groups (44% as in the entire trial), nearly 100 low-risk individuals with a 10-year risk of cardiovascular events of 5% would need to be treated for 5 years to prevent 1 event. On the other hand, <25 high-risk individuals (10-year risk 20%) would need to be treated for 5 years to prevent an event. Thus, reclassifying individuals as low or high risk could have immediate clinical relevance with regard to identifying candidates for statin therapy, from the standpoint of absolute risk reduction. This is illustrated by the results of a recent post hoc analysis of the JUPITER data, which showed that individuals with higher CRP levels had greater absolute benefit from rosuvastatin (due to their higher baseline risk), even though the relative risk reductions were somewhat lower. Thus, by demonstrating the efficacy of statin treatment across a broad range of LDL and risk factor levels, the JUPITER results highlight the value of any test that allows refinement of absolute risk estimation.

Cost Effectiveness

Even if randomized trial data support the clinical effectiveness of a biomarker-guided strategy, it must be determined whether such a strategy can be implemented at a reasonable cost. One of the questions raised after the JUPITER trial was whether the modest reductions in the absolute risk of cardiovascular events justify the costs of widespread CRP testing combined with statin therapy. Furthermore, other “costs” should be considered, including the potential harm from long-term statin use, such as the risk of diabetes mellitus, which might not be fully captured in short-term trials. Serious consideration of any preventive strategy should include a detailed cost-benefit analysis that integrates multiple inputs: performance characteristics of the screening test (typically from observational studies), effectiveness of the biomarker-guided therapy (randomized trials), incidence of side effects (trials and postmarketing studies), and financial cost (observational and administrative data). Because it is impractical to perform randomized trials for every potential biomarker, the results of such cost-benefit analyses (with the use of a range of assumptions for therapeutic effectiveness) could be used to help prioritize among competing biomarkers.

With any biomarker-guided strategy, expensive and/or high-risk interventions make good risk stratification more valuable because it is important to avoid exposing low-risk individuals to the therapy. Conversely, with inexpensive and safe therapies, risk stratification becomes less valuable. Indeed, because current preventive therapies such as statins or aspirin are generally regarded as well tolerated, some have questioned the need for any risk-based screening at all, especially in older adults. “Polypill” strategies have been
proposed as an alternative to risk-based screening.119 The Indian Polycap Study (TIPS) showed that a combination pill containing aspirin, simvastatin, ramipril, hydrochlorothiazide, and atenolol was well tolerated in a relatively unsenescing, middle-aged to elderly population.120 The long-term outcomes of the polypill approach are presently unknown.121 Even if the polypill is shown to improve outcomes with an acceptable safety profile, risk prediction is likely to remain important for risk communication, motivating adherence and lifestyle modification, and selecting candidates for more aggressive (“personalized”) therapies.121,122

## Conclusion

The dramatic increase in the number of biomarker and imaging studies in the past decade has heightened the debate over whether such novel tests are ready for routine clinical use. This tension is perhaps most evident in the area of primary prevention, where there is a large opportunity to reduce the burden of cardiovascular disease but with the costs and potential risks inherent in screening large populations. The scientific, political, and regulatory implications of this debate are further evidenced by recent developments surrounding the dissemination of cardiac imaging and the measurement of CRP for targeting statin therapy.

A number of circulating, genetic, and imaging biomarkers are robustly associated with cardiovascular risk. Nonetheless, the evidence that these biomarkers enhance risk prediction for individuals is surprisingly limited (as summarized in Table 2). Even in combination, most biomarkers (with the possible exception of coronary calcium scores) produce minimal improvements in discrimination and calibration. Although some data suggest that biomarkers can reclassify individuals within risk categories, this benefit is largely confined to those at intermediate risk of cardiac events. The value of reclassification for altering clinical management in this group remains largely theoretical because randomized trials specifically evaluating this question have not been performed.

Despite the limitations of current cardiovascular biomarkers, it is worth recognizing the tremendous progress in the field that has taken place over the past decade, spanning experimental studies, epidemiological investigations, and clinical trials. Furthermore, the pace of biomarker discoveries is likely to accelerate in coming years, with the maturation of technologies such as proteomics, metabolomics, and transcriptomics. Such approaches may be particularly useful for identifying biomarkers in different pathways from those represented by existing biomarkers. Coupling these advances with a better understanding of which subsets of the population require further risk stratification and the initiation of randomized trials to test these hypotheses could make biomarker-guided prevention an attainable goal.

## Acknowledgments

The author gratefully acknowledges Dr Michael Pencina, Boston University, for his permission to include the simulation results shown in Figure 2.

## Sources of Funding

Dr Wang was supported in part by National Institutes of Health grants R01-HL-083197, R01-HL-083875, R01-DK-081572, and R01-HL098283 and a grant from the American Heart Association.

## Disclosures

Dr Wang participates in the scientific advisory board of Diasorin Inc. He is named as a coinventor on patent applications for the use of proadrenomedullin in risk stratification. He has been a coauthor in prior studies that received support for the measurement and the use of proadrenomedullin in risk stratification. He has been a coauthor in prior studies that received support for the measurement of biomarkers from Siemens Healthcare Diagnostics, Brahms, Critical Diagnostics, and Singulex.

## References


### Table 2. Summary of Evidence for Selected Biomarkers in Risk Prediction

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Association With Outcomes</th>
<th>Improves Discrimination</th>
<th>Improves Calibration</th>
<th>Reclassifies a Large Proportion</th>
<th>Reclassification Is Accurate</th>
<th>Result Should Affect Clinical Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRP</td>
<td>Yes</td>
<td>No</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>Unknown</td>
</tr>
<tr>
<td>BNP</td>
<td>Yes</td>
<td>+/-/-†</td>
<td>No</td>
<td>+/-</td>
<td>+/-</td>
<td>Unknown</td>
</tr>
<tr>
<td>SNPs at 9p21</td>
<td>Yes</td>
<td>No†</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Coronary artery calcium score</td>
<td>Yes</td>
<td>Yes</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yes</td>
<td>Unknown</td>
</tr>
<tr>
<td>Carotid IMT</td>
<td>Yes</td>
<td>No†</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Pulse wave velocity</td>
<td>Yes</td>
<td>+/-/-†</td>
<td>Unknown</td>
<td>+/-</td>
<td>Yes</td>
<td>Unknown</td>
</tr>
<tr>
<td>Ankle-brachial index</td>
<td>Yes</td>
<td>+/-/-‡</td>
<td>Unknown</td>
<td>+/-</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

SNP indicates single-nucleotide polymorphisms. For references, see text. Modified from May and Wang,53 with permission of the publisher. Copyright © 2008, Elsevier.

*Some evidence in high-risk populations.
†Based on a small number of studies.
‡Some evidence in women.


Assessing the Role of Circulating, Genetic, and Imaging Biomarkers in Cardiovascular Risk Prediction
Thomas J. Wang

Circulation. 2011;123:551-565
doi: 10.1161/CIRCULATIONAHA.109.912568

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/123/5/551

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

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

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/