The MESA heart failure risk score: can’t we do more?

Jennifer E Ho, Jared W Magnani

In 1987, Framingham Heart Study pioneer Dr William B Kannel wrote, ‘Epidemiologic data on incidence and prognosis of cardiac failure in the general population are actually quite sparse.’ Dr Kannel proceeded to report risk factors for heart failure identified in Framingham that included hypertension, ECG LV hypertrophy, obesity, diabetes, radiographic cardiac enlargement and cigarette use. Over the last decades heart failure epidemiology has been informed by advances in cardiac imaging, identification of novel biomarkers and refinements in pathophysiology and classification.

The exposures listed by Dr Kannel remain salient to heart failure risk. However, many individuals develop cardiovascular disease even in their absence. Risk assessment is refined by the integration of imaging, biomarkers and novel assessments. Risk scores are fundamental and essential for advancing risk prediction, and serve multiple functions to enhance epidemiological and clinical assessment. First, risk scores provide an avenue to integrate established exposures with novel, contemporary assessments in risk quantification. Second, risk scores may target at-risk populations and refine clinical definitions—such as ‘stage A’ and ‘stage B’ heart failure, for example—with the goal of disease prevention. Third, individualised risk scores can provide personal assessments of risk, be a tool for patient education or focus efforts to optimise prevention. Ultimately, the utility of a risk score is determined by its clinical relevance: can it be employed to target preventive strategies in heart failure, and ‘turn back the clock’ for a disease where the median survival upon diagnosis is as dismal as 5 years? Is it relevant to diverse populations? Can the addition of novel imaging or biomarkers be implemented in a cost-effective manner?

To our knowledge, three major, observational, community-based cohort studies have published heart failure risk scores. In 1999, Kannel et al described the probability of heart failure in Framingham participants with coronary disease, hypertension or valvular heart disease. The study did not include biomarkers or imaging aside from the chest radiograph, limiting its contemporary applicability, and preceded the advent of classification and discrimination, recognised now as essential for critically evaluating risk functions. The Framingham score was tested in the Dynamics of Health, Aging and Body Composition (Health ABC) study. Strengths of the model included the cohort’s biracial design and enrolment of older adults, a population at increased heart failure risk. It is no surprise that the Framingham risk model had limited risk discrimination (C-statistic <0.70) in Health ABC; risk models translate poorly across cohorts with different designs, covariate measurement, and event ascertainment and adjudication. More recently, Atherosclerosis Risk in Communities (ARIC) Study investigators derived an ARIC heart failure risk model and tested the Framingham and Health ABC functions. The results are telling: first, the heart failure risk score derived in ARIC was very robust (C-statistic=0.80). Second, the Framingham and Health ABC scores performed better with estimates derived in the ARIC cohort, rather than those published with the original cohort data. Third, all three risk scores improved in ARIC with the inclusion of N-terminal B-type natriuretic peptide (BNP). A consistent lesson is that scores perform better in their derivation cohort. The selected characteristics of the heart failure risk scores described are summarised in table 1.

The most recent articulation of a heart failure risk is presented by Chahal et al. They present a novel score developed from the Multi-Ethnic Study of Atherosclerosis (MESA) Study in over 6600 MESA participants. The score includes readily accessible covariates that have survived application in the other heart failure risk scores. MESA’s ethnic and racial diversity enhances this heart failure score. We commend Chahal et al for a cogent discussion that situates MESA’s ethnic/racial composition in the context of the other risk scores cited here.

A fundamental strength of the presentation by Chahal et al consists in presenting nested models that integrate biomarker and imaging assessments. In separate models, the investigators determined the relative contributions of BNP and LV mass index (LVMI) as quantified by cardiac magnetic resonance. Presenting the comprehensive data with and without these assessments is noteworthy; such an approach provides a transparent evaluation of how assessments bolster risk prediction. The addition of BNP improved the C-statistic from 0.80 for the baseline model to 0.87 with its inclusion. Net reclassification improvement (NRI) was similarly enhanced (0.37, CIs not provided in manuscript). Interestingly, while the model was strengthened with the addition of LVMI, adding LVMI on top of BNP yielded only modest improvement. However, adding BNP to a model with LVMI yielded a 15% NRI (CIs not provided). The take-home is that BNP emerged a critical and salient contributor towards heart failure risk prediction.

We would like to suggest several limitations with the presentation by Chahal et al. The follow-up duration is shorter than that of the other scores described here, likely limiting the number of identified cases. The Framingham study employed cross-sectional pooling to evaluate 4-year risk windows, leveraging decades of follow-up. The ARIC score identified a heart failure incidence of 11% during a 15.5-year follow-up. In contrast, the present study identified a 3% event incidence. Heart failure develops insidiously, so we expect more cases will be identified prospectively as subclinical disease becomes more manifest in the MESA cohort. An argument could have been made for pursuing another iteration of this project as the cohort ages. Second, the investigators present race-specific estimates of discrimination. However, the absence of NRI, likely because of the small number of events, renders the data difficult to interpret. Hence, the race-specific generalisability of the MESA heart failure risk score, as the authors acknowledge, is limited. Third, individuals with prevalent cardiovascular disease were excluded from MESA, limiting generalisability to those with established increased risk for developing heart failure. In contrast, the Framingham and Health ABC risk scores were developed in higher-risk cohorts. We suggest that the approaches are complementary; risk score development across varied cohorts improves our understanding of heart failure risk. A final limitation
<table>
<thead>
<tr>
<th>Cohort</th>
<th>Year published</th>
<th>Heart failure definition</th>
<th>Risk-window (years)</th>
<th>Risk factors included in base model</th>
<th>Summary findings</th>
<th>Major strengths</th>
<th>Selected limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framingham Heart Study</td>
<td>1999</td>
<td>Symptoms, physical exam features</td>
<td>4</td>
<td>Age, ECG LVH, heart rate, systolic BP, CHD, valvular disease, diabetes, with and without radiologic cardiomegaly and vital capacity</td>
<td>Publication of regression coefficients Articulation of risk profile with point scoring</td>
<td>Broad age range (45–94 years) Long-term follow-up (38 years) Pooled regression model Heart failure adjudication</td>
<td>Racially homogenous cohort Absence of biomarker or imaging assessments No discrimination or reclassification metrics Generalisability (CHD, hypertension, valvular heart disease)</td>
</tr>
<tr>
<td>Health ABC</td>
<td>2008</td>
<td>Overnight hospitalisation</td>
<td>5</td>
<td>Age, CHD, ECG LVH, heart rate, systolic BP, albumin, creatine, glucose, smoking</td>
<td>C-statistic 0.73 Framingham Heart Study score C-statistic &lt;0.70</td>
<td>Older adults Bicardal (black and white race) Internal validation</td>
<td>Heart failure determination by hospitalisation only Misclassification of baseline covariates Cohort size (n=2853) Generalisability (cohort is black and white race)</td>
</tr>
<tr>
<td>Atherosclerosis Risk in Communities Study</td>
<td>2012</td>
<td>ICD-9 coding, hospitalisation or death certificates</td>
<td>10</td>
<td>Age, sex, race, heart rate, systolic BP, CHD, antihypertensive medication, diabetes, BMI, smoking, log NT-BNP</td>
<td>C-statistic 0.80 (without NT-BNP) C-statistic 0.81 (with inclusion of NT-BNP)</td>
<td>Inclusion of BNP Bicardal (black and white race) Long-term follow-up Covariate range</td>
<td>Heart failure ascertainment by hospitalisation, death certificate Generalisability (cohort is black and white race)</td>
</tr>
<tr>
<td>Multi-Ethnic Study of Atherosclerosis</td>
<td>2014</td>
<td>Self-report and medical record</td>
<td>5</td>
<td>Age, sex, race and ethnicity, heart rate, systolic BP, diabetes, BMI, smoking, cholesterol, creatine, with and without log NT-BNP and LVMI</td>
<td>C-statistic 0.80 (without NT-BNP) C-statistic 0.87 (with inclusion of NT-BNP) Significant reclassification (25%–37%) with addition of LVMI or BNP</td>
<td>Multithread design Inclusion of BNP and LVMI Covariate range</td>
<td>Follow-up duration (median 4.7 years) with low heart failure incidence Younger-aged cohort Limited events, precluding race-specific risk score development</td>
</tr>
</tbody>
</table>

BMI, body mass index; BP, blood pressure; CHD, coronary heart disease; Health ABC, Dynamics of Health, Aging and Body Composition; ICD-9, International Classification of Diseases, Ninth Revision; LVH, LV hypertrophy; LVMI, LV mass index; NT-BNP, N-terminal B-type natriuretic peptide.
evaluate their contributions. Third, the MESA heart failure risk score is the fourth such score to our knowledge developed in a community-based cohort. We would argue that the time has come for cardiovascular risk prediction to move beyond the individual cohort. A well-designed multicohort heart failure risk score would have increased generalisability and facilitated conducting the race-specific and ethnic-specific approaches that could not be undertaken in the MESA analysis. In conclusion, the current era has seen epidemiology face increasing emphasis on how it can modify public health. Risk scores are informative at the cohort level, and the study by Chahal et al is a substantive contribution to this literature. We now need to demonstrate that the heart failure risk score can be applied towards disease prevention.

Contributors JEH and JWM participated in the drafting of this editorial and take full responsibility for its content.

Competing interests This work was supported by grants from the National Institute of Health to JWM (R03AG045075) and to JEH (K23-HL116780). JWM and JEH are further supported by Boston University School of Medicine Department of Medicine Career Investment Awards.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Ho JE, Magnani JW. Heart 2015;101:7–9.

Published Online First 11 November 2014

http://dx.doi.org/10.1136/heartjnl-2014-305697

REFERENCES


Editorial

Ho JE, et al. Heart January 2015 Vol 101 No 1