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

Jennifer E Ho,1,2 Jared W Magnani1,2

In 1987, Framingham Heart Study pioneer Dr William B Kannel1 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?2 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 al3 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.4 The Framingham score was tested in the Dynamics of Health, Aging and Body Composition (Health ABC) study.5 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.6 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.7 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 be 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

1Department of National Heart, Lung and Blood Institute’s and Boston University’s Framingham Heart Study, Framingham, Massachusetts, USA;2Section of Cardiovascular Medicine, Department of Medicine, Boston University, Boston, Massachusetts, USA

Correspondence to Dr Jared W Magnani, Section of Cardiovascular Medicine, Boston University School of Medicine, 88 E. Newton Street, Boston 02118, MA, USA; jmagmani@bu.edu

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

<table>
<thead>
<tr>
<th>Year</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>1999</td>
<td>Epidemiological, physical examination features</td>
<td>4</td>
<td>Age, sex, physical examination features, alcohol consumption, smoking</td>
<td>C-statistic 0.73</td>
<td>Framingham Heart Study score</td>
<td>Generalisability (CHD, hypertension, valvular heart disease)</td>
</tr>
<tr>
<td>2008</td>
<td>Overnight hospitalisation</td>
<td>5</td>
<td>Age, sex, physical examination features, alcohol consumption, smoking, BMI, blood pressure, diabetes, cholesterol, uric acid</td>
<td>C-statistic 0.80 (without NT-BNP)</td>
<td>C-statistic 0.81 with inclusion of NT-BNP (U)</td>
<td>Generalisability (heart failure risk score)</td>
</tr>
<tr>
<td>2012</td>
<td>ICD-9 coding, hospitalisation or death certificates</td>
<td>10</td>
<td>Age, sex, heart rate, systolic BP, diastolic BP, diabetes, peripheral arterial disease, heart failure, smoking, log NT-BNP</td>
<td>C-statistic 0.87 (without NT-BNP)</td>
<td>C-statistic 0.97 with inclusion of NT-BNP (U)</td>
<td>Generalisability (heart failure risk score)</td>
</tr>
<tr>
<td>2014</td>
<td>Self-report and medical record</td>
<td>5</td>
<td>Age, sex, race and ethnicity, heart rate, systolic BP, diastolic BP, diabetes, BMI, smoking, cholesterol, glucose, creatinine, with and without log NT-BNP and UMI</td>
<td>C-statistic 0.80 (without NT-BNP)</td>
<td>C-statistic 0.87 with inclusion of NT-BNP (U)</td>
<td>Generalisability (heart failure risk score)</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; U, UMI, UMI, LI, LV mass index.

is the absence of external validation. The C-statistic of the final risk score was exceptionally high. It would be important to validate the performance of the MESA model in an external population. Cross-cohort development and validation of a heart failure risk score will enhance the impact and relevance of the product. Lastly, the NRI is perhaps less meaningful, since no universally accepted categories exist with respect to heart failure risk, and thus no change in treatment is prompted. Further, the use of four risk categories can increase reclassification, thereby inflating the NRI, again highlighting the importance of external validation. Furthermore, a ‘very high’ risk category, defined as >20% heart failure risk, is likely superfluous.

More than an academic exercise, deriving generalisable, novel risk scores has significant potential to enhance prevention. A primary objective of the analysis presented by the MESA investigators is a model informed by variables available in the primary care setting. The authors describe the tool as applicable towards motivating ‘both patients and physicians’ to target modifiable risk factors. We applaud the implication of a partnership between the patient and physician to modify and address risk.

We are far from done with heart failure risk prediction. The fundamental question still remains: what is the clinical utility of a heart failure risk score? We can learn from the application of risk scores to cardiovascular disease prevention in general and their utility in identifying the at-risk patient. Can a similar preventive strategy be guided by a heart failure risk score? The St Vincent’s Screening to Prevent Heart Failure Study recently randomised 1374 patients free of heart failure to usual care versus screening with BNP testing and showed reduced rates of LV systolic and diastolic dysfunction and clinical heart failure after a 4.2-year follow-up.

We believe that similar future studies are needed to test the direct clinical relevance of risk scores to disease prevention. Future studies of heart failure risk may also consider distinguishing separate risks for preserved and reduced ejection fraction. The presentations and disease course differ between the two and merit the development of specific risk functions. Second, multiple biomarkers have been associated with heart failure. High-sensitivity troponin, soluble-ST2 and galectin-3 are examples. As such markers become more mainstream, their inclusion will be important even in the ‘parsimonious model’, as the MESA authors describe their heart failure risk score, in order to...
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 (1R03AG045075) and to JEH (K23-HL116780). JWM and JEH are further supported by Boston University School of Medicine Department of Medicine Career Development 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

REFERENCES


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

Jennifer E Ho and Jared W Magnani

*Heart* 2015 101: 7-9 originally published online November 11, 2014
doi: 10.1136/heartjnl-2014-306459

Updated information and services can be found at:
http://heart.bmj.com/content/101/1/7

References

This article cites 8 articles, 5 of which you can access for free at:
http://heart.bmj.com/content/101/1/7#BIBL

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

- Drugs: cardiovascular system (8669)
- Epidemiology (3660)
- Health policy (221)
- Hypertension (2935)
- Clinical diagnostic tests (4710)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/