Critical Review Form
Clinical Prediction or Decision Rule

Can the HEART score safely reduce stress testing and cardiac imaging in patients at low risk for major adverse cardiac events? Crit Pathw Cardiol. 2011 Sep;10(3):128-33.

Objectives: "to determine whether the HEART score is predictive of MACE and to evaluate its potential for safely reducing objective cardiac testing in a United States ED cohort of low-risk chest pain patients." (p. 128)

Methods: This was a retrospective cohort study including patients enrolled in the ED-based observation unit (OU) Chest Pain Registry at Wake Forest Baptist Medical Centre between January 2008 and April 2010. The registry consisted of patients with chest pain concerning for acute coronary syndrome (ACS) determined to be low risk (TIMI Score < 2 and clinician assessment of low risk) admitted to the OU with a normal or nondiagnostic ECG and a negative first set of cardiac biomarkers (troponin I, CK-MB).

The primary outcome was a composite of major adverse cardiovascular events (MACE) including all-cause mortality, myocardial infarction (defined by a troponin cut-off of 1.0 ng/mL), or coronary revascularization during the index visit or within 30 days (defined as angioplasty, coronary stent placement, or coronary artery bypass grafting). The predictor variable was a modified HEART score (Table 1) that did not include BMI. A normal troponin value of 0.065 ng/mL was used. Components of the score were determined by chart review. Patients with missing or incomplete data underwent a blinded chart review by study investigators using a standardized data collection tool. Data abstraction was duplicated on 50 patients to assess for interobserver agreement. A HEART Score of 0-3 was considered "low risk" while a score of ≥ 4 was considered "high risk." A secondary analysis was conducted using a combination of the HEART Score and a second serum troponin drawn 4-6 hours after ED admission.

During the study period, 1070 low-risk chest pain patients were included in the OU Chest Pain Registry. The mean age was 46.3, 56.5% were white, 38.8% were African American, and 60.6% were male. Of the total cohort, 532 required a chart review for incomplete data. Interobserver agreement among the blinded investigators for low versus high-risk HEART Scores was calculated (kappa = 0.81). Stress testing or cardiac imagining occurred in 1003 (93.7%) patients. Among patients without MACE at the index visit (n = 1058), complete 30-day follow-up data, based on chart review, was available for 753 (70%) patients.

Table 1. The HEART Score	
	Points
History • **Highly suspicious** • **Moderate suspicious** • **Slightly suspicious**	 **2** **1** **0**
ECG • **Significant ST-depression** • **Nonspecific repolarization abnormality** • **Normal**	 **2** **1** **0**
Age • **≥ 65** • **45-65** • **≤ 45**	 **2** **1** **0**
Risk factors • **3 or more risk factors** • **1-2 risk factors** • **No risk factors**	 **2** **1** **0**
Troponin • **≥ 3 × normal limit** • **1-3 × normal limit** • **≤ normal limit**	 **2** **1** **0**
Risk factors include diabetes mellitus, smoker, hypertension, hypercholesterolemia, family history of coronary artery disease, obesity, or a history of significant atherosclerosis (coronary revascularization, myocardial infarction, stroke, or peripheral arterial disease).	

Guide		Comments
I.	***Is this a newly derived instrument*** **(Level IV)?**	
A.	Was validation restricted to the retrospective use of statistical techniques on the original database? (If so, this is a Level IV rule & is not ready for clinical application).	No. This was a retrospective review of a new cohort of patients, different from the original database.
II.	**Has the instrument been validated? (Level II or III). If so, consider the following:**	
1a	Were all important predictors included in the derivation process?	N/A
1b	Were all important predictors present in significant proportion of the study population?	N/A
1c	Does the rule make clinical sense?	Yes. The HEART Score includes a combination of known historical risk factors, ECG findings, age, cardiac enzymes values, and clinical suspicion.

2	Did validation include prospective studies on several different populations from that used to derive it (II) or was it restricted to a single population (III)?	Validation of the CDR in this study was limited to the retrospective application of the rule to a single cohort of patients at a single study site.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	Yes. While the study was limited to low-risk chest pain patients admitted to the OU (TIMI score < 2 and clinician assessment of low risk), there was a wide range of HEART scores (0-6), suspicion of ACS, age, and number of risk factors. While there were no patients with ST depression, and all patients had troponin levels below the upper limit of normal, one would expect this in a low-risk population. The MACE rate varies from 0.6% in those with a low-risk HEART score to 4.2% in those with a high-risk HEART score.
3b	Was there a blinded assessment of the gold standard?	Uncertain. There was no single gold-standard test in this study. The outcome of interest was a composite outcome, including mortality, myocardial infarction, and need for revascularization. Assessment for the outcome was performed "a record review" to determine the incidence of MACE within 30 days. The authors do not explicitly state if those performing the record review were blinded to HEART Scores or to study purpose.
3c	Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?	Uncertain. Assessment of the predictor variables was performed by review of registry data, although the authors do not explicitly state if the data abstractors were blinded to outcomes. In patients with incomplete data in the registry, a blinded chart review was conducted to gather additional information
3d	Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?	No. There was no gold standard, but rather the outcome of interest was a composite MACE outcome. All patients in the study underwent record review to determine the presence or absence of outcomes. The decision to perform stress testing on patients (which could influence the decision to perform revascularization) was not influenced by the HEART Score itself, as this was a retrospective review, but the individual components of the score likely did influence this decision.
4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or	MACE occurred in 1.1% (12/1070) of patients in the cohort. Among patients with a low-risk HEART Score, 0.6% (5/904) (95% CI 0.2–1.1%) had a MACE.

	absolute outcome rates)?	Among patients with a high-risk HEART Score 4.2% (7/166) (95% CI: 1.9–8.6%) had a MACE: (OR 7.92; 95% CI: 2.48–25.25; $P < 0.001$). • Sensitivity: 58.3% (95% CI 32-81%) • Specificity: 85% (95% CI 83-87%) • LR+: 3.89 (2.36-6.39) • LR-: 0.49 (95% CI: 0.25-0.96) Based on these results, use of the HEART score to determine the need for stress testing would have resulted in 5 cases of missed ACS, a miss rate of less than 0.5% (5/1070; 95% CI 0.2–1.1%) and reduction in cardiac testing of 84.5% (904/1070; 95% CI: 82–86.5%). Combing the HEART Score with a 4-6 hour troponin level resulted in 100% sensitivity (95% CI 72-100%) and specificity of 83.1% (95% CI 81-85%), with a reduction in cardiac stress testing of 82.1% (95% CI 80-84%).
III.	**Has an impact analysis demonstrated change in clinical behavior or patient outcomes as a result of using the instrument? (Level I). If so, consider the following:**	
1	How well did the study guard against bias in terms of differences at the start (concealed randomization, adjustment in analysis) or as the study proceeded (blinding, co-intervention, loss to follow-up)?	Poorly. This was a retrospective analysis on data that was partly collected prospectively and partly retrospectively. The authors removed one of the components of the HEART Score (BMI) as this information was not readily available from the chart. They did not mention if they blinded outcome assessors to the results of the CDR, and did not describe well the methods of assessing for the components of the outcome. They state simply that a record review was performed to assess for the outcome, but do not state if a standardized form was used to accomplish this, and do not assess the interobserver reliability of the process. In addition, 305 (28.5%) patients were did not have complete follow-up data, and in the analysis these patients were considered not to have had any of the adverse outcomes. There was no sensitivity analysis on this patient population. Additionally, all 12 adverse events occurred on the index visit, making the accuracy of follow-up on the remaining patients questionable.
2	What was the impact on clinician behavior and patient-important outcomes?	This was a retrospective study and the results of the CDR were not used to impact clinical outcomes. The authors concluded that use of the HEART Score had

		the potential to reduce stress testing in 84.5% of patients, and the combination of the HEART Score and a 4-6 hour troponin had the potential to reduce stress testing in 82.1% of patients.

Limitations:

1. **Given that this was a retrospective chart review, it is unclear how the "Level of suspicion" was determined (as a component of the HEART score).**

2. **One of the requirements for inclusion was reportedly "clinician assessment of low risk", and yet 53 patients (5%) were designated "highly suspicious" on the HEART score.**

3. **The method of patient follow-up was not well defined, and appears to involve chart review at the investigating site with no telephone follow-up or record review from surrounding institutions.**

4. **The chart review methods were not well-described (Gilbert 1996 and Worster 2004).**

5. **28.5% of patients did not have complete 30-day follow-up data. All 12 adverse outcomes occurred on the index visit, raising concern that adverse events occurring after the index visit may have been missed on follow-up.**

6. **MI was defined by troponin elevation alone (level > 1.0 ng/mL).**

7. **There is no true "gold standard" in the assessment of outcomes in chest pain patients. Instead, a composite outcome (MACE) is used whose components are not necessarily equivalent in terms of patient importance. This practice has been called into question (Kip 2008).**

Bottom Line:

This study demonstrated that use of a HEART score of 0-3 in an already low-risk population of chest pain patients (TIMI < 2, normal or non-diagnostic ECG, negative initial set of cardiac enzymes) predicted a very low risk of major adverse cardiac events (5/904, 0.6%). Unfortunately, this was a retrospective study with poorly defined follow-up methodology and a large number of patients with incomplete follow-up data (28.5%). The requirement that the TIMI score be used prior to calculating the HEART score would also likely make this protocol difficult to employ.