HEART Score Risk Stratification of Low-Risk Chest Pain Patients in the Emergency Department: A Systematic Review and Meta-Analysis

Jessica Laureano-Phillips, MPH; Richard D. Robinson, MD; Subhash Aryal, PhD; Somer Blair, PhD; Damalia Wilson, MD; Kellie Boyd, MLS; Chet D. Schrader, MD; Nestor R. Zenarosa, MD; Hao Wang, MD, PhD*

*Corresponding Author. E-mail: hwang01@jpshealth.org.

Study objective: The objectives of this systematic review and meta-analysis are to appraise the evidence in regard to the diagnostic accuracy of a low-risk History, ECG, Age, Risk Factors, and Troponin (HEART) score for prediction of major adverse cardiac events in emergency department (ED) patients. These included 4 subgroup analyses: by geographic region, the use of a modified low-risk HEART score (traditional HEART score [0 to 3] in addition to negative troponin results), using conventional versus high-sensitivity troponin assays in the HEART score, and a comparison of different post–ED-discharge patient follow-up intervals.

Methods: We searched MEDLINE, EBSCO, Web of Science, and Cochrane Database for studies on the diagnostic performance of low-risk HEART scores to predict major adverse cardiac events among ED chest pain patients. Two reviewers independently screened articles for inclusion, assessed the quality of studies with both an adapted Quality Assessment of Diagnostic Accuracy Studies version 2 tool and an internally developed tool that combined components of the Quality in Prognostic Studies; Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies; and Grading of Recommendations Assessment, Development and Evaluation. Pooled sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios were calculated.

Results: There were 25 studies published from 2010 to 2017, with a total of 25,266 patients included in the final meta-analysis, of whom 9,919 (39.3%) were deemed to have low-risk HEART scores (0 to 3). Among patients with low-risk HEART scores, short-term major adverse cardiac events (30 days to 6 weeks) occurred in 2.1% of the population (182/8,832) compared with 21.9% of patients (3,290/15,038) with non–low-risk HEART scores (4 to 10). For patients with HEART scores of 0 to 3, the pooled sensitivity of short-term major adverse cardiac event predictions was 0.96 (95% confidence interval [CI] 0.93 to 0.98), specificity was 0.42 (95% CI 0.36 to 0.49), positive predictive value was 0.19 (95% CI 0.14 to 0.24), negative predictive value was 0.99 (95% CI 0.98 to 0.99), positive likelihood ratio was 1.66 (95% CI 1.50 to 1.85), and negative likelihood ratio was 0.09 (95% CI 0.06 to 0.15). Subgroup analysis showed that lower short-term major adverse cardiac events occurred among North American patients (0.7%), occurred when modified low-risk HEART score was used (0.8%), or occurred when high-sensitivity troponin was used for low-risk HEART score calculations (0.8%).

Conclusion: In this meta-analysis, despite its use in different patient populations, the troponin type used, and timeline of follow-up, a low-risk HEART score had high sensitivity, negative predictive value, and negative likelihood ratio for predicting short-term major adverse cardiac events, although risk of bias and statistical heterogeneity were high. [Ann Emerg Med. 2019;74:187-203.]

Please see page 188 for the Editor's Capsule Summary of this article.

A **podcast** for this article is available at www.annemergmed.com.

0196-0644/$-see front matter

https://doi.org/10.1016/j.annemergmed.2018.12.010

INTRODUCTION

Background

Chest pain accounts for approximately 5% to 10% of US emergency department (ED) visits and is one of the most common complaints among all ED visits worldwide.[1,2] A significant number of chest pain patient visits result in hospitalization to rule out acute coronary syndrome, although few patients require cardiac interventions.[3-5] This clinical practice consequently results in substantial patient overassessment, wasted allocation of already limited health care resources, and increased financial burden for both patients and the overall health care system.

In recent years, multiple chest pain risk-stratification scoring systems (eg, Thrombolysis in Myocardial Infarction, Global Registry of Acute Coronary Events) have

Editor's Capsule Summary

What is already known on this topic

The History, ECG, Age, Risk Factors, and Troponin (HEART) score has been shown to have high sensitivity in many emergency department–based studies around the world, but surveys of practicing physicians suggest it needs to be even higher (≥98%) to be used clinically.

What question this study addressed

This study confirmed the findings of a previous systematic review and also examined prespecified subgroups for whom the sensitivity of a low-risk HEART score might be higher.

What this study adds to our knowledge

After removal of studies that excluded patients with any positive troponin results, the studies that tested the modified HEART score and those that used only high-sensitivity troponins had the highest sensitivity for predicting short-term major adverse cardiac events (97% and 98%, respectively). Lower 95% confidence intervals were 96% and 91%, respectively.

How this is relevant to clinical practice

For the HEART score to be used clinically, these results suggest that future work should target the modified HEART score, high-sensitivity troponins, or both.

been derived and validated in different studies.[6-8] However, most scoring systems are intended to evaluate short-term outcomes, such as mortality or recurrence of myocardial infarction, among hospitalized chest pain patients with history of coronary artery disease. Although their use can be expanded to all chest pain patients, scant evidence exists to support such claims.[9,10] Originally derived in Europe, the History, ECG, Age, Risk Factors, and Troponin (HEART) score is an alternative chest pain risk-stratification tool intended specifically to identify low-risk chest pain patients at the ED and predict potential major adverse cardiac events after ED discharge.[11,12] Chest pain patients with low HEART scores (0 to 3) might be safely discharged from the ED, with very low risk of major adverse cardiac events, in effect potentially minimizing unnecessary hospital admissions and thereby considerably reducing health care costs. To date, one systematic review and meta-analysis has been published on the use of the HEART score in the management of ED chest pain patients with suspected acute coronary syndrome.[13]

Importance

Although the HEART score has been broadly used and validated in clinical studies, it has not been formally included in the American Heart Association guidelines.[14,15] A recent meta-analysis showed that 1.6% of patients with low-risk HEART scores would have had a missed major adverse cardiac event at a mean follow-up time of 6 weeks.[13] This pooled major adverse cardiac event occurrence among low-risk chest pain patients still exceeds the acceptable less than 1% miss rate among ED providers willing to discharge these patients directly from the ED.[16] Many potential factors could affect the diagnostic and prognostic accuracy of HEART score predictions, including differences in patient populations, provider determination of low-risk HEART score criteria, specific troponin reagent used, and follow-up duration, resulting in high clinical heterogeneity.

Goals of This Investigation

We aimed to find an optimal use of the HEART score that may provide improved major adverse cardiac event predictions and a more meaningful guideline for ED providers, thus decreasing the rate of post–ED-discharge major adverse cardiac events in low-risk chest pain patients. We reviewed all relevant studies to assess the diagnostic performance of the low-risk HEART score as a predictor of major adverse cardiac events among ED chest pain patients receiving care in different geographic regions, using different low-risk HEART score criteria, using different troponin reagents, and at distinct post–ED-discharge follow-up intervals.

MATERIALS AND METHODS

Study Design and Selection of Participants

A dedicated medical librarian searched MEDLINE (including PubMed and Ovid), EBSCO (Cumulative Index of Nursing and Allied Health, including "choose/select" all databases feature), Web of Science, and the Cochrane Database for publications of the HEART score as a diagnostic factor to predict short-term (within 6 weeks) and long-term (3 months or greater) major adverse cardiac event outcomes. The literature search was filtered to include studies published from 2008 (the year in which the HEART score was initially derived) through May 2017 inclusive.[11] The patient population was limited to those treated in the ED or ED observation unit and included dedicated chest pain observation units. Additional searches

were conducted with Scopus and Google Scholar, and an exhaustive manual search of relevant publication reference lists was conducted (Figure E1, available online at http://www.annemergmed.com).

Eligible studies included original prospective and retrospective cohort studies and randomized controlled trials in peer-reviewed journals in either print or online ahead-of-print publications and those available in open-access journals. Additional study eligibility criteria consisted of adult patients (≥18 years) who initially presented at the ED or were ultimately assigned to the ED observation unit (acute chest pain unit) with acute nontraumatic chest pain or its equivalent, with potential risk of acute coronary syndrome; HEART score calculation, either prospectively or retrospectively; and complete postdischarge follow-up data for the occurrence of major adverse cardiac events. Given that our analysis focused mainly on diagnostic performance of the low-risk HEART score to predict all-timeframe major adverse cardiac event outcomes among ED chest pain patients, studies were excluded according to the following criteria: non-English-language publications; conference abstracts and posters, commentaries, case reports, editorials, and reviews; non-ED or non-ED observation unit patient populations; noncardiac chest pain with definitive diagnoses during the index visit; secondary data analysis with duplicated data; no clear cardiac outcomes addressed nor predictive outcomes measured; and studies addressing only inhospital major adverse cardiac event outcomes without any follow-up data from discharged patients.

Two independent reviewers (J.L.-P. and D.W.) screened search results in 3 phases: title, abstract, and full text. Study selection progressed from one phase to the next if both reviewers agreed to include the individual study. When disagreements occurred, a third reviewer (H.W.) was consulted to determine the individual study's eligibility. The same method was used for study quality assessment. If data were missing, incomplete, or deemed insufficient to calculate sensitivity, specificity, negative predictive value, positive predictive value, and positive and negative likelihood ratios, at least 2 e-mails were sent to the study's corresponding author to request raw study data. When a study cohort was described in more than one study, we included only the article best meeting the inclusion criteria. Studies were excluded according to quality assessment if a cohort was accounted for in another included study, no outcomes data of interest were reported, missing or unclear data were discerned, a published protocol was without data, the study setting was not an ED or ED observation unit, or HEART score was not calculated (Figure 1). Reviewers (J.L.-P. and D.W.) independently extracted data into a predesigned data collection form. The following key information was abstracted: first author; publication year; study patient population; sample size; study location (eg, ED, ED observation unit); study inclusion/exclusion criteria; patient sex and race; HEART score calculations, including the methods of calculations; low-risk HEART score definitions; type of troponin reagent used; different outcomes measurements; and follow-up periods.

Possible points for each of the 5 components of the HEART score are 0 to 2, resulting in a total HEART score range of 0 to 10 (Table E1, available online at http://www.annemergmed.com).[12] To maximize the number of eligible studies for the meta-analysis, minor differences in each study's HEART score definition were allowed. The detailed differences of HEART score components pertaining to defining chest pain suspicious for ischemia, how ECGs were evaluated, and the inclusion of obesity as a risk factor are listed in Table E2, available online at http://www.annemergmed.com. For troponin, we included both conventional and high-sensitivity troponin. Some studies[17,18] did not address troponin reagent specifically, whereas others[19-21] used a mix of conventional and high-sensitivity troponin reagents. Additionally, a modified low-risk HEART score was used in several studies, in which only patients with negative troponin values were considered low risk despite having a total HEART score less than or equal to 3. Such studies were analyzed separately in this meta-analysis.[14,21-28] Thus patients with a chronically elevated troponin (e.g. patients with renal failure, in whom the repeat troponin is the same) were excluded.

Outcome Measures

The primary endpoint was occurrence of major adverse cardiac events from 30 days to 6 weeks. A major adverse cardiac event was traditionally defined as acute myocardial infarction, coronary revascularization by percutaneous coronary intervention with or without additional interventions, coronary artery bypass graft surgery, and all-cause mortality.[12] Minor outcome measurement deviations included non–ST-elevation myocardial infarction, fatal or nonfatal acute myocardial infarction,[19] ventricular arrhythmia or high-degree atrioventricular block requiring intervention, cardiac death, and life-threatening cardiac complications.[29] Various myocardial infarction definitions were used among the included studies and comprised annotations according to the European Society of Cardiology, American College of Cardiology, American Heart Association, and Work Heart Federation guideline.[30] This meta-analysis included major adverse cardiac events and other outcome measurements during specific follow-up intervals. For this study, we divided the assorted follow-up

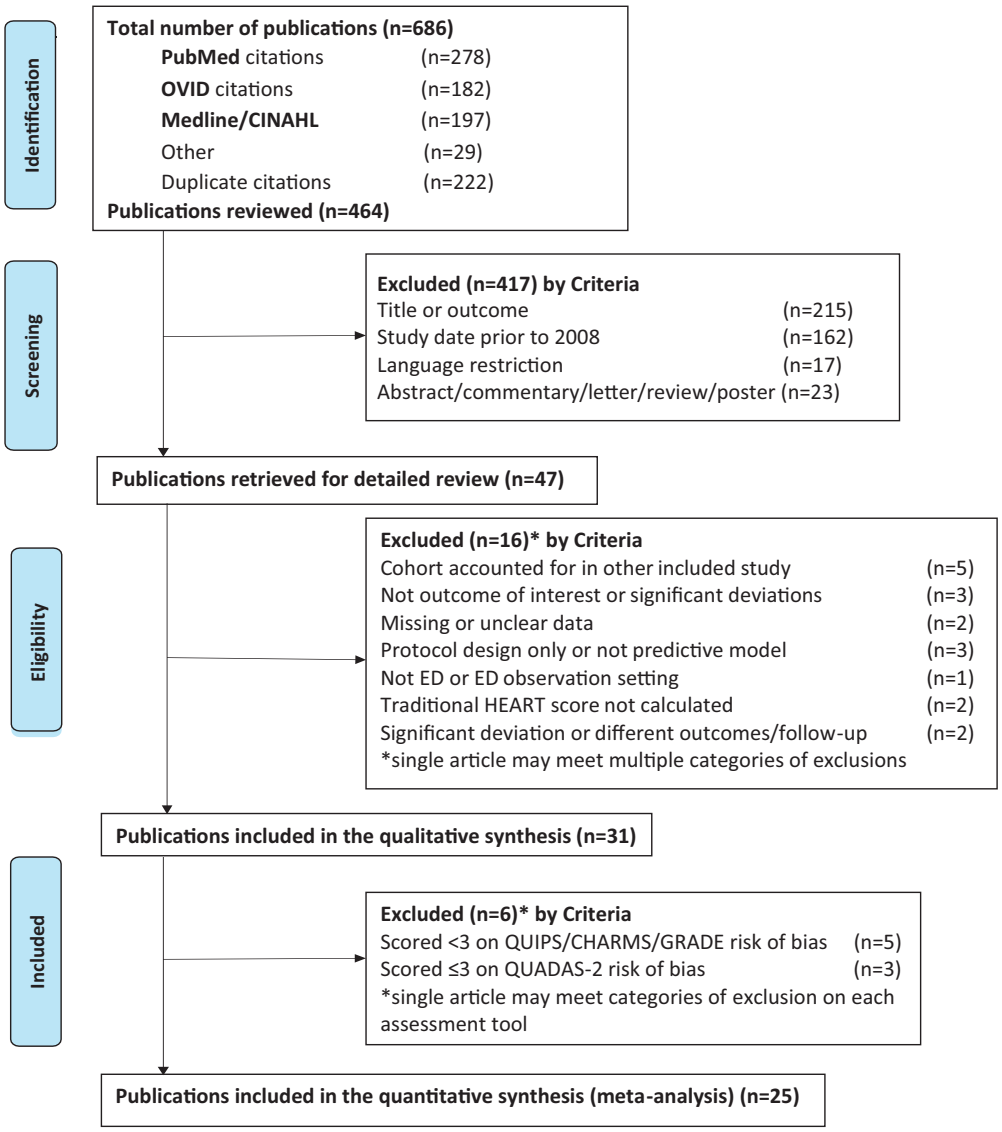

Figure 1. Included studies' search strategy flowchart.

intervals into either short term (30 days to 6 weeks) or long term (3 months or greater).

We exercised dual screening tools for individual, study-level risk of bias and quality assessment (a dual risk-of-bias assessment approach was necessary because the HEART score is considered both a diagnostic and prognostic tool). Our initial assessment tool combined elements of the Quality in Prognostic Studies; the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies; and the Grading of Recommendations Assessment, Development and Evaluation. In general, the Quality in Prognostic Studies is broadly used for quality assessment in prognostic studies,[31] the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies provides reporting guidance for the systematic review of prognostic studies,[32] and the Grading of Recommendations Assessment, Development and Evaluation is intended for the assessment of randomized controlled trials.[33] Although components of each tool overlap, all are designed to judge the quality of evidence when prognostic research studies are reviewed. Briefly, this amended tool included 5 categories: source of data and participation, study attrition, prognostic factor measurement, outcome measurement, and statistical analysis and reporting. Reviewers scored each of the 5 risk-of-bias categories as low, uncertain, or high. Low risk scored 1 point, uncertain risk scored 0.5 point, and high risk scored 0 points. A possible total score ranged from 0 to 5 points. Studies with a total score of less than 3 were considered to have a high risk of bias and were therefore

excluded from the final meta-analysis. The second risk-of-bias assessment used was an adapted Quality Assessment of Diagnostic Accuracy Studies version 2 tool.[34] Again, the same scoring system was used to assess the risk of bias for each study, with a total score ranging from 0 to 7 points. A study was excluded from quantitative analysis if it scored 3 or less on the adapted assessment tool.

The assessment of potential meta-biases, such as publication bias and small-study effects, is required for meta-analyses. Publication bias is often examined by checking for asymmetry in funnel plots. However, the literature is inconsistent and the ability to identify publication bias with funnel plots is essentially equal to chance (53% accuracy).[35] We performed both Harbord and Peters tests with a modified Galbraith plot because both tests maintain better control of the false-positive rate than the traditional Egger's test for small-study effects.[36-39] To account for selective publication bias, it is prudent to determine what an unbiased data set might contribute and reanalyze the overall effect by adding any apparent "missing" studies. Accordingly, we performed a publication bias analysis by using the Duval and Tweedie[40] nonparametric "trim and fill" method. This method, which was developed to estimate the number of missing studies that might exist in a meta-analysis and the effect that they might have on the study's outcome, was used to further examine publication bias.[40] Such an analytic approach enhances the accuracy of evaluating publication bias and minimizes inconsistency across different assessments.

Primary Data Analysis

For quality assessment of individual studies, we used κ statistics for interrater variability assessment. Individual study population data were pooled to assess the sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of the HEART score for the prediction of all-timeframe major adverse cardiac events. Given the expectation of study heterogeneity, a random-effects model was used. Cochrane's Q statistic, with the null hypothesis assuming that all studies in the analysis shared a common effect size, and I^2 statistics, providing the proportion of the observed variance reflecting differences in the true effect sizes rather than sampling error, were used for heterogeneity analysis. The hierarchic summary receiver operating characteristic (ROC) was used for meta-analysis of HEART score accuracy in a logistic regression model.[41]

Additionally, different heterogeneities occurred among systematic reviews and meta-analysis studies, including both clinical heterogeneity and methodological or statistical heterogeneity. Clinical heterogeneity refers to differences in clinically related characteristics that can give rise to variations in pooled treatment-effects estimates.[42] Two clinicians (D.W. and H.W.) estimated the study cohorts' clinical heterogeneity caused by patient characteristics, clinical variables, and different timeframe major adverse cardiac event outcomes. If discrepancies occurred, a third reviewer provided the final decision. We determined 4 areas that could potentially affect clinical heterogeneity. Therefore, 4 subgroup analyses were performed: a comparison of diverse patient populations as presented from 3 geographic regions (North America, Europe, and Asia-Pacific), a comparison of traditional versus modified low-risk HEART score criteria, a comparison of conventional-troponin and high-sensitivity-troponin assays on that component of the HEART score, and a comparison of short- versus long-term major adverse cardiac events. Their pooled sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio were also reported. The meta-analysis was conducted with Stata (version 14.2; StataCorp, College Station, TX).

This meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses[43] and Meta-analysis of Observational Studies in Epidemiology reporting guidelines.[44]

RESULTS

Characteristics of Study Subjects

Our literature search identified a total of 686 potentially relevant studies. After further independent review of titles and abstracts, 47 studies were considered promising for inclusion. After full-text review and quality assessments, 25 studies met inclusion criteria for this meta-analysis. Details concerning the exclusion process are shown in Figure 1. A summary of all 47 reviewed studies is shown in Table E3, available online at http://www.annemergmed.com.

Main Results

A total of 25 studies were assessed in this meta-analysis, comprising data from 25,266 patients (range 94 to 3,456 patients), among whom 9,919 (39.3%) were deemed to have low-risk HEART scores. Characteristics of the 25 enrolled studies are listed in Tables 1 and 2. Of the 25 included studies, in regard to major adverse cardiac events, 20 reported short-term occurrence,[12,14,17-20,22,24-26,28,45-53] 3 reported both short- and long-term occurrence,[18,54,55] and only 2 showed long-term occurrence.[23,27]

The occurrence of short-term major adverse cardiac events among all enrolled studies was 14.5% (3,472/23,870) (Table 3). In patients with low-risk HEART scores

Table 1. Summary of enrolled studies included in the systemic review and meta-analysis.

First Author	Year	Study Design	No. of Sites	Sample Size (n)	Men/Women, No. (%)	Race and Ethnicity, No. (%)	Outcomes (MACE and Others*)
Backus[12]	2010	Retrospective	4	880	521 (57)/359 (43)	No race/ethnicity frequencies provided	MACE
Mahler[24]	2011	Retrospective	1	1,070	648 (61)/422 (40)	White, 605 (57) Black, 415 (39) Asian, 13 (1) Other, 37 (4)	MACE
Backus[45]	2013	Prospective	10	2,388	1,372 (58)/1,016 (43)	No race/ethnicity frequencies provided	MACE and others
Mahler[25]	2013	Prospective	18	1,005	545 (54)/460 (46)	White, 676 (67) Black, 256 (25) Hispanic, 55 (6) Other, 18 (2)	MACE and others
Melki[27]	2013	Retrospective	1	410	222 (54)/188 (46)	Swedish, 410 (100)	MACE
Six[46]	2013	Prospective	14	2,906	1,754 (60)/1,152 (40)	Chinese, 660 (23) White, 1,709 (59)	MACE
Carlton[19]	2015	Prospective	1	959	564 (59)/395 (41)	White British, 913 (95) (Only frequencies for British patients provided)	Fatal and nonfatal AMI
Leite[47]	2015	Retrospective	1	174	129 (74)/45 (26)	No race/ethnicity frequencies provided	MACE
Mahler[14]	2015	Prospective	1	282	120 (43)/162 (57)	White, 183 (65) Black, 94 (33) Asian, 1 (0.3) Native American, 2 (0.7) Others, 2 (0.7) **Ethnicity** Hispanic, 5 (2) Non-Hispanic, 277 (98)	MACE
Visser[48]	2015	Prospective	1	255	142 (56)/113 (44)	No race/ethnicity frequencies provided	MACE and others
Baugh[26]	2016	Prospective	1	97	No frequencies for sex	No race/ethnicity frequencies provided	MACE and others
Bodapati[49]	2016	Retrospective	1	678	368 (54)/310 (46)	Indigenous Australian, 148 (22) Non-Indigenous Australian, 530 (78)	MACE
Bolvardi[17]	2016	Prospective	1	100	57 (57)/43 (43)	No race/ethnicity frequencies provided	MACE and others
Chen[54]	2016	Prospective	2	833	461 (55)/372 (45)	Chinese, 833 (100)	MACE and others
Jain[50]	2016	Retrospective	1	968	427 (45)/520 (55) (Only 947 patients evaluated for outcome analysis)	White, 149 (16) Black, 771 (81) Other, 27 (3)	MACE
Rainer[51]	2016	Prospective	1	602	294 (49)/308 (51)	Chinese, 595 (99)	MACE

Sakamoto[52]	2016	Retrospective	1	604	418 (69)/186 (31)	Chinese, 368 (61) Malay, 115 (19) Indian, 97 (16) Other, 24 (4)	MACE
Santi[55]	2016	Retrospective	1	1,378	778 (57)/600 (44)	No race/ethnicity frequencies provided	MACE and others
Wang[23]	2016	Prospective	1	986	539 (55)/447 (45)	White, 374 (38) Black, 341 (35) Other, 271 (27) **Ethnicity** Hispanic, 256 (26) Non-Hispanic, 727 (74) Unknown, 3 (0.3)	MACE
Dubin[28]	2017	Retrospective	1	771	347 (45)/424 (55)	Black, 449 (83) White, 53 (10) Hispanic/Latino, 25 (5) Hawaiian, 2 (0.0) (Race recorded for 538 patients only)	MACE
Bank[53]	2017	Prospective	1	1,915	1,084 (57)/831 (43)	No race/ethnicity frequencies provided	MACE and others
de Hoog[20]	2017	Prospective	3	3,456	2,128 (62)/1,328 (38)	White, 1,791 (52) Chinese, 1,059 (31) Indian, 344 (10) Malay, 262 (8)	MACE
McCord[22]	2017	Prospective	12	661	385 (58)/276 (42)	White, 551 (83)	MACE
Patnaik[18]	2017	Retrospective	1	299	152 (51)/147 (49)	White, 23 (8) Black, 245 (82) Asian, 5 (2) Hispanic, 21 (7) Other, 5 (2)	MACE and others
Poldervaart[21]	2017	Prospective	9	3,648	1,980 (54)/1,668 (46)	No race/ethnicity frequencies provided	MACE and others

MACE, Major adverse cardiac event; *AMI*, acute myocardial infarction.

*Mace and others: MACE includes acute myocardial infarction (fatal or nonfatal), coronary angiography (emergent/elective with or without stent placement), coronary artery bypass grafting, and all-cause death. Others include unstable angina, acute coronary syndrome, coronary angiography revealing significant stenosis but with conservative therapy (the presence of known significant coronary stenosis is thought to be the cause of chest pain without revascularization because of co-morbidity or high risk of complications, patient refusal or resolution of symptoms after determination of a contributory cause, such as tachydysrhythmia), cardiac arrest, cardiogenic shock, ventricular arrhythmia requiring intervention, high-degree atrioventricular block requiring intervention, or ventricular arrhythmia requiring intervention (including percutaneous radiofrequency ablation and pacemaker implantation).

Table 2. Summary of diagnostic accuracy of low-risk HEART scores predictive of major adverse cardiac event outcomes among enrolled studies.

First Author	Year	Location	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR– (95% CI)	Low-Risk Criteria*	Low Risk, %	MACE in Low Risk, %	Follow-up Period†	Troponin Type‡
Backus[12]	2010	EU	0.98 (0.95–1.00)	0.42 (0.38–0.45)	0.27 (0.23–0.31)	0.99 (0.97–1.00)	1.68 (1.57–1.79)	0.05 (0.01–0.14)	Traditional	34.4 (303/880)	1.0 (3/303)	6 wk	Trop I/T
Mahler[24]	2011	NA	0.58 (0.28–0.85)	0.85 (0.83–0.87)	0.04 (0.02–0.08)	0.99 (0.99–1.00)	3.88 (2.36–6.39)	0.49 (0.25–0.96)	Modified	84.5 (904/1,070)	0.6 (6/904)	30 days	Trop I
Backus[45]	2013	EU	0.96 (0.94–0.98)	0.43 (0.41–0.45)	0.26 (0.24–0.28)	0.98 (0.97–0.99)	1.69 (1.62–1.77)	0.09 (0.05–0.14)	Traditional	36.4 (870/2,388)	1.7 (15/870)	6 wk	Trop I/T
Mahler[25]	2013	NA	0.99 (0.96–1.00)	0.28 (0.24–0.31)	0.26 (0.23–0.29)	0.99 (0.97–1.00)	1.37 (1.31–1.43)	0.04 (0.01–0.14)	Modified	22.2 (220/991)	0.9 (2/220)	30 days	Trop I
Melki[27]	2013	EU	0.97 (0.83–1.00)	0.65 (0.60–0.70)	0.18 (0.12–0.25)	1.00 (0.98–1.00)	2.74 (2.36–3.19)	0.05 (0.01–0.35)	Traditional	60.2 (247/410)	0.4 (1/247)	3 mo	Trop I/T
Melki[27]	2013	EU	0.92 (0.64–1.00)	0.66 (0.61–0.71)	0.09 (0.05–0.15)	1.00 (0.98–1.00)	2.75 (2.23–3.41)	0.12 (0.02–0.76)	Modified	64.5 (245/380)	0.4 (1/245)	3 mo	Trop I/T
Six[46]	2013	AP	0.96 (0.94–0.98)	0.32 (0.30–0.34)	0.17 (0.16–0.19)	0.98 (0.97–0.99)	1.41 (1.37–1.46)	0.12 (0.07–0.20)	Traditional	28.2 (820/2,906)	1.7 (14/820)	30 days	Trop I/T
Carlton[19]	2015	EU	0.94 (0.86–0.98)	0.34 (0.31–0.37)	0.11 (0.09–0.14)	0.98 (0.96–0.99)	1.42 (1.31–1.53)	0.19 (0.08–0.44)	Traditional	31.6 (303/959)	1.7 (5/303)	30 days	Hs-Trop I/T
Leite[47]	2015	EU	0.90 (0.70–0.99)	0.63 (0.55–0.70)	0.25 (0.16–0.36)	0.98 (0.93–1.00)	2.43 (1.90–3.11)	0.15 (0.04–0.57)	Traditional	56.3 (98/174)	2.0 (2/98)	6 wk	Trop I
Mahler[14]	2015	NA	1.00 (0.63–1.00)	0.50 (0.41–0.58)	0.11 (0.05–0.20)	1.00 (0.95–1.00)	1.87 (1.49–2.36)	0.11 (0.01–1.00)	Modified	46.8 (66/141)	0 (0/66)	30 days	Trop I
Visser[48]	2015	EU	0.93 (0.85–0.98)	0.44 (0.37–0.52)	0.41 (0.34–0.49)	0.94 (0.87–0.98)	1.68 (1.45–1.94)	0.15 (0.06–0.36)	Traditional	33.3 (85/255)	5.9 (5/85)	6 wk	Hs-Trop T
Baugh[26]	2016	NA	1.00 (0.40–1.00)	0.60 (0.49–0.70)	0.10 (0.03–0.24)	1.00 (0.93–1.00)	2.24 (1.53–3.30)	0.17 (0.01–1.00)	Modified	57.4 (54/94)	0 (0/54)	30 days	Trop T
Bodapati[49]	2016	AP	0.99 (0.95–1.00)	0.43 (0.39–0.48)	0.32 (0.27–0.36)	0.99 (0.97–1.00)	1.74 (1.62–1.88)	0.03 (0.01–0.13)	Traditional	34.7 (235/678)	0.9 (2/235)	30 days	Trop I
Bolvardi[17]	2016	AP	1.00 (0.86–1.00)	0.14 (0.07–0.24)	0.27 (0.18–0.37)	1.00 (0.72–1.00)	1.15 (1.03–1.28)	0.13 (0.01–1.00)	Traditional	11.0 (11/100)	0 (0/11)	30 days	N/A
Chen[54]	2016	AP	0.88 (0.79–0.94)	0.41 (0.37–0.45)	0.15 (0.12–0.19)	0.97 (0.94–0.98)	1.49 (1.35–1.64)	0.30 (0.17–0.52)	Traditional	37.8 (315/833)	3.5 (11/315)	30 days	Trop T
Chen[54]	2016	AP	0.89 (0.82–0.94)	0.42 (0.39–0.46)	0.21 (0.17–0.25)	0.96 (0.93–0.98)	1.55 (1.42–1.69)	0.25 (0.15–0.43)	Traditional	37.8 (315/833)	4.1 (13/315)	6 mo	Trop T
Jain[50]	2016	NA	0.99 (0.96–1.00)	0.21 (0.19–0.24)	0.17 (0.15–0.20)	0.99 (0.97–1.00)	1.26 (1.22–1.31)	0.03 (0.01–0.24)	Traditional	18.5 (175/947)	0.6 (1/175)	30 days	Trop I
Rainer[51]	2016	AP	0.98 (0.87–1.00)	0.40 (0.36–0.44)	0.11 (0.08–0.15)	1.00 (0.98–1.00)	1.63 (1.50–1.77)	0.06 (0.01–0.41)	Traditional	37.5 (226/602)	0.4 (1/226)	30 days	Hs-Trop T

Sakamoto[52]	2016	AP	0.99 (0.97–1.00)	0.25 (0.21–0.30)	0.42 (0.38–0.47)	0.98 (0.93–1.00)	1.32 (1.24–1.40)	0.04 (0.01–0.15)	Traditional	16.4 (99/604)	2.0 (2/99)	30 days	Trop T
Santi[55]	2016	EU	1.00 (0.98–1.00)	0.44 (0.41–0.47)	0.24 (0.21–0.27)	1.00 (0.99–1.00)	1.77 (1.68–1.86)	0.01 (0.01–0.09)	Traditional	37.2 (512/1,378)	0 (0/512)	30 days	Hs-Trop T
Santi[55]	2016	EU	1.00 (0.98–1.00)	0.45 (0.42–0.48)	0.27 (0.24–0.30)	1.00 (0.99–1.00)	1.80 (1.71–1.90)	0.00 (0.01–0.08)	Traditional	37.2 (512/1,378)	0 (0/512)	6 mo	Hs-Trop T
Wang[23]	2016	NA	0.57 (0.29–0.82)	0.86 (0.83–0.88)	0.05 (0.02–0.11)	0.99 (0.98–1.00)	4.02 (2.49–6.50)	0.50 (0.27–0.92)	Modified	85.2 (840/986)	0.7 (6/840)	6 mo	Trop I
Dubin[28]	2017	NA	0.91 (0.81–0.97)	0.32 (0.28–0.35)	0.10 (0.07–0.13)	0.98 (0.95–0.99)	1.34 (1.22–1.47)	0.27 (0.12–0.63)	Traditional	30.0 (232/773)	2.2 (5/232)	6 wk	Trop I
Dubin[28]	2017	NA	0.95 (0.86–0.99)	0.31 (0.27–0.34)	0.10 (0.08–0.13)	0.99 (0.96–1.00)	1.37 (1.27–1.48)	0.17 (0.06–0.51)	Modified	28.7 (222/773)	1.4 (3/222)	6 wk	Trop I
Bank[53]	2017	EU	0.90 (0.86–0.93)	0.42 (0.39–0.44)	0.23 (0.21–0.25)	0.96 (0.94–0.97)	1.54 (1.46–1.63)	0.24 (0.17–0.34)	Traditional	36.7 (702/1,915)	4.4 (31/702)	6 wk	Trop I
de Hoog[20]	2017	EU/AP	0.88 (0.85–0.91)	0.45 (0.43–0.47)	0.22 (0.21–0.24)	0.95 (0.94–0.96)	1.60 (1.53–1.67)	0.27 (0.21–0.34)	Traditional	39.8 (1,376/3,456)	4.6 (63/1,376)	6 wk	Hs-Trop T, Trop I
McCord[22]	2017	NA/EU/AP	0.83 (0.36–1.00)	0.63 (0.59–0.67)	0.02 (0.01–0.05)	1.00 (0.99–1.00)	2.25 (1.55–3.26)	0.26 (0.04–1.00)	Modified	62.5 (413/661)	0.2 (1/413)	30 days	Hs-Trop T
Patnaik[18]	2017	NA	1.00 (0.93–1.00)	0.40 (0.34–0.46)	0.26 (0.20–0.33)	1.00 (0.96–1.00)	1.64 (1.48–1.82)	0.02 (0.01–0.38)	Traditional	32.8 (98/299)	0 (0/98)	6 wk	N/A
Patnaik[18]	2017	NA	0.96 (0.89–0.99)	0.43 (0.36–0.50)	0.37 (0.30–0.44)	0.97 (0.91–0.99)	1.68 (1.49–1.90)	0.09 (0.03–0.28)	Traditional	32.8 (98/299)	3.1 (3/98)	1 y	N/A
Poldervaart[21]	2017	EU	0.96 (0.93–0.98)	0.49 (0.46–0.51)	0.30 (0.27–0.33)	0.98 (0.97–0.99)	1.87 (1.77–1.98)	0.09 (0.05–0.15)	Traditional	40.5 (715/1,766)	2.0 (14/715)	6 wk	Hs-Trop T; Trop I,
Poldervaart[21]	2017	EU	0.97 (0.94–0.99)	0.49 (0.46–0.51)	0.30 (0.27–0.33)	0.99 (0.97–0.99)	1.89 (1.79–2.00)	0.06 (0.03–0.11)	Modified	40.3 (711/1,766)	1.4 (10/711)	6 wk	Hs-Trop T; Trop I

PPV, Positive predictive value; *NPV*, negative predictive value; *LR+*, positive likelihood ratio; LR–, negative likelihood ratio; *EU*, Europe; *Trop*, troponin; *NA*, North America; *AP*, Asia-Pacific; *Hs-Trop*, high-sensitivity troponin; *N/A*, not applicable.

*Traditional refers to conventional HEART score calculations, with HEART score 0 to 3 being low risk; modified low-risk criteria refer to using a traditional HEART score of 0 to 3, in addition to negative troponin results.

†Refers to short term (either 30 days or 6 weeks) or long term (longer than 6 weeks; eg, 3 months, 6 months, or 1 year).

‡Refers to either conventional troponin T or I or Hs-Trop T or I.

Table 3. A comparison of major adverse cardiac event occurrence among subgroups.

	Occurrence of MACEs, %	
Overall and Subgroup	**Low-Risk Group**	**High-Risk Group**
Overall short-term MACE*	2.1 (182/8,832)	21.9 (3,290/15,038)
Overall long-term MACE*	1.1 (23/2,012)	23.7 (449/1,894)
Overall all-timeframe MACE*	1.9 (189/9,919)	21.7 (3,327/15,347)
Subgroup analysis		
Geographic locations		
North America (short term)	0.7 (13/1,745)	17.8 (456/2,566)
North America (all-timeframe MACE)	0.7 (19/2,585)	17.1 (464/2,712)
Europe (short-term MACE)	2.5 (105/4,230)	24.3 (1,771/7,276)
Europe (all-timeframe MACE)	2.4 (106/4,477)	24.2 (1,800/7,439)
Asia-Pacific (short-term and all-timeframe MACE)	2.6 (63/2,440)	21.4 (1,058/4,948)
HEART criteria		
Traditional HEART criteria (short-term MACE)	2.4 (174/7,175)	22.3 (3,068/13,738)
Traditional HEART criteria (all-timeframe MACE)	2.4 (175/7,422)	22.3 (3,097/13,901)
Modified HEART criteria (short-term MACE)	0.8 (21/2,590)	20.5 (596/2,906)
Modified HEART criteria (all-timeframe MACE)	0.8 (28/3,675)	19.7 (633/3,217)
Troponin reagent		
Conventional troponin (short-term MACE)	1.8 (93/5,093)	21.9 (2,039/9,301)
Conventional troponin (all-timeframe MACE)	1.6 (100/6,180)	21.6 (2,076/9,610)
High-sensitivity troponin (short-term and all-timeframe MACE)	0.8 (12/1,539)	17.1 (396/2,316)
Mixed troponin† (short-term and all-timeframe MACE)	3.5 (77/2,200)	25.0 (855/3,421)

*Short term refers to patients with either 30-day or 6-week follow-up after the index ED visits; long term refers to at least more than 3-month follow-up after the index ED visits. All timeframe is defined as all 23 short-term MACE studies, in addition to 2 long-term ones (Melki[27] and Wang[23]).
†These studies used mixed troponin reagents for the HEART calculation, which occurred in multicenter studies, and were unable to be further divided into either conventional or high-sensitivity troponin groups, or the studies did not specify the type of troponin reagents used.

(0 to 3), occurrence of short-term major adverse cardiac events was 2.1% (182/8,832) compared with 21.9% (3,290/15,038) among those with non–low-risk HEART scores (4 to 10). Pooled sensitivity of short-term major adverse cardiac event predictions among low-risk HEART scores was 0.96 (95% confidence interval [CI] 0.93 to 0.98), specificity was 0.42 (95% CI 0.36 to 0.49), positive predictive value was 0.19 (95% CI 0.14 to 0.24), negative predictive value was 0.99 (95% CI 0.98 to 0.99), positive likelihood ratio was 1.66 (95% CI 1.50 to 1.85), and negative likelihood ratio was 0.09 (95% CI 0.06 to 0.15) (Table 4). A forest plot of pooled sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio for prediction of short-term major adverse cardiac events among low-risk HEART scores is shown in Figure 2. *Q* statistics of sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio all reached statistical significance ($P<.001$), indicating significantly high heterogeneity between individual study effects and the pooled effect across studies. I^2 statistics measurements were greater than 75, indicating that the percentages of variation across studies were due to high heterogeneity rather than chance.

The occurrence of all-timeframe and long-term major adverse cardiac events among enrolled studies was 13.9% (3,516/25,266) and 12.1% (472/3,906), respectively. In patients with low-risk HEART scores (0 to 3), the occurrence of all-timeframe and long-term major adverse cardiac events was 1.9% (189/9,919) and 1.1% (23/2,012), respectively. Major adverse cardiac event occurrences and pooled sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of subgroups are reported in Tables 3 and 4 and further addressed in detail under subgroup analysis.

Furthermore, hierarchic summary ROC and the area under the curve of the low-risk HEART scores predictive of short-term and all-timeframe major adverse cardiac events were 0.78 (95% CI 0.74 to 0.81) and 0.80 (95% CI 0.77 to 0.84), respectively, indicating fair discrimination (Figure E2, available online at http://www.annemergmed.com).

Subgroup comparisons of the occurrence of short-term and all-timeframe major adverse cardiac events, including

Table 4. A comparison of the diagnostic accuracy of low-risk HEART score predictive of major adverse cardiac event outcomes.

	Pooled Sensitivity (95% CI)	Pooled Specificity (95% CI)	Pooled PPV (95% CI)	Pooled NPV (95% CI)	Pooled LR+ (95% CI)	Pooled LR– (95% CI)
Overall short-term MACE*	0.96 (0.93–0.98)	0.42 (0.36–0.49)	0.19 (0.14–0.24)	0.99 (0.98–0.99)	1.66 (1.50–1.85)	0.09 (0.06–0.15)
Overall long-term MACE*	0.96 (0.76–0.99)	0.59 (0.41–0.74)	0.21 (0.11–0.35)	0.99 (0.97–1.00)	2.31 (1.60–3.34)	0.08 (0.01–0.40)
Overall all-timeframe MACE*	0.96 (0.93–0.98)	0.45 (0.38–0.53)	0.18 (0.14–0.23)	0.99 (0.98–0.99)	1.74 (1.54–1.98)	0.10 (0.06–0.15)
Subgroup meta-analysis						
Geographic locations						
North America (short term MACE)	0.96 (0.87–0.99)	0.45 (0.29–0.62)	0.13 (0.08–0.21)	0.99 (0.98–1.00)	1.75 (1.32–2.34)	0.08 (0.03–0.25)
North America (all-timeframe MACE)	0.94 (0.82–0.98)	0.51 (0.33–0.70)	0.12 (0.08–0.19)	0.99 (0.99–1.00)	1.94 (1.38–2.72)	0.11 (0.04–0.29)
Europe (short-term MACE)	0.96 (0.92–0.98)	0.44 (0.40–0.48)	0.25 (0.20–0.30)	0.98 (0.96–0.99)	1.71 (1.57–1.85)	0.10 (0.06–0.18)
Europe (all-timeframe MACE)	0.96 (0.92–0.98)	0.46 (0.41–0.52)	0.24 (0.20–0.29)	0.98 (0.97–0.99)	1.78 (1.59–1.99)	0.09 (0.05–0.17)
Asia-Pacific (short-term and all-timeframe MACE)	0.97 (0.91–0.99)	0.35 (0.27–0.43)	0.22 (0.16–0.30)	0.98 (0.96–0.99)	1.48 (1.34–1.65)	0.09 (0.04–0.21)
HEART criteria						
Traditional HEART criteria (short-term MACE)	0.97 (0.94–0.98)	0.38 (0.33–0.43)	0.22 (0.18–0.27)	0.98 (0.97–0.99)	1.56 (1.45–1.68)	0.08 (0.05–0.14)
Traditional HEART criteria (all-timeframe MACE)	0.97 (0.94–0.98)	0.39 (0.34–0.45)	0.22 (0.18–0.26)	0.98 (0.98–0.99)	1.60 (1.47–1.74)	0.08 (0.05–0.14)
Modified HEART criteria (short-term MACE)	0.95 (0.88–0.98)	0.53 (0.37–0.68)	0.10 (0.05–0.19)	0.99 (0.99–1.00)	2.02 (1.51–2.69)	0.09 (0.05–0.19)
Modified HEART criteria (all-timeframe MACE)	0.93 (0.84–0.97)	0.59 (0.44–0.72)	0.10 (0.06–0.17)	0.99 (0.99–1.00)	2.27 (1.69–3.05)	0.11 (0.06–0.22)
Troponin reagent						
Conventional troponin (short-term MACE)	0.96 (0.91–0.98)	0.43 (0.34–0.53)	0.19 (0.14–0.25)	0.99 (0.98–0.99)	1.67 (1.44–1.94)	0.10 (0.06–0.18)
Conventional troponin (all-timeframe MACE)	0.95 (0.90–0.97)	0.48 (0.37–0.58)	0.18 (0.13–0.23)	0.99 (0.98–0.99)	1.81 (1.51–2.17)	0.11 (0.07–0.19)
High-sensitivity troponin (short-term and all-timeframe MACE)	0.97 (0.88–0.99)	0.45 (0.36–0.54)	0.13 (0.05–0.29)	0.99 (0.97–1.00)	1.76 (1.51–2.06)	0.07 (0.02–0.26)
Mixed troponin† (short-term and all-timeframe MACE)	0.99 (0.35–1.00)	0.36 (0.24–0.51)	0.26 (0.23–0.30)	0.97 (0.95–0.99)	1.56 (1.27–1.91)	0.02 (0.00–4.02)

*Short term refers to patients with either 30-day or 6-week follow-up after the index ED visits, long-term refers to at least more than 3-month follow-up after the index ED visits, and all timeframe is defined as analyzing data from all 25 enrolled studies, including all 23 short-term MACE studies and 2 long-term ones.

†These studies used mixed troponin reagents for the HEART calculation, which occurred in multicenter studies, and they were unable to be further divided into either conventional or high-sensitivity troponin groups.

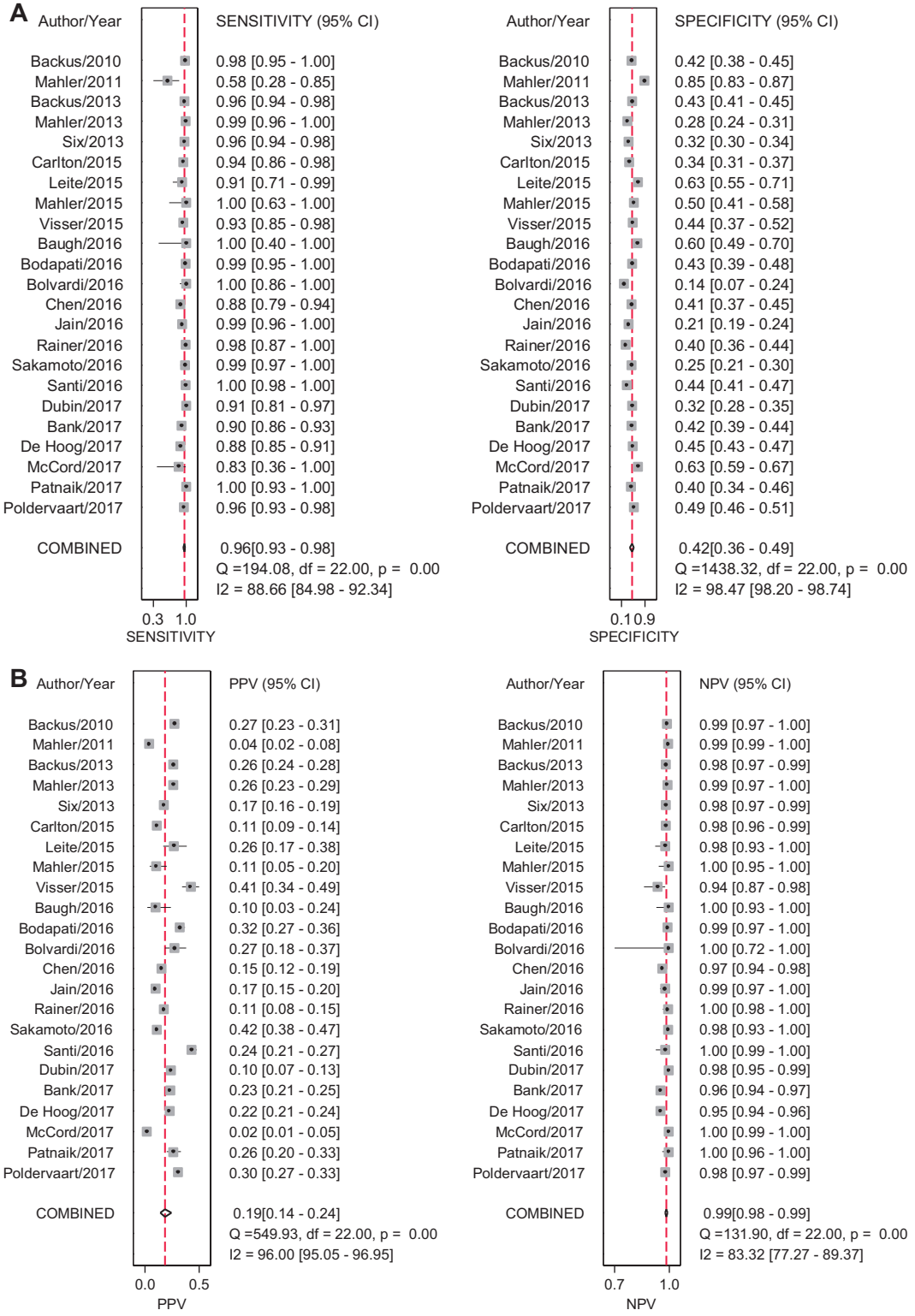

Figure 2. *A*, Pooled sensitivity and specificity of low-risk HEART scores predictive of short-term major adverse cardiac event outcomes. *B*, Pooled positive and negative predictive values of low-risk HEART scores predictive of short-term major adverse cardiac event outcomes. *C*, Pooled positive and negative likelihood ratios of low-risk HEART scores predictive of short-term major adverse cardiac event outcomes.

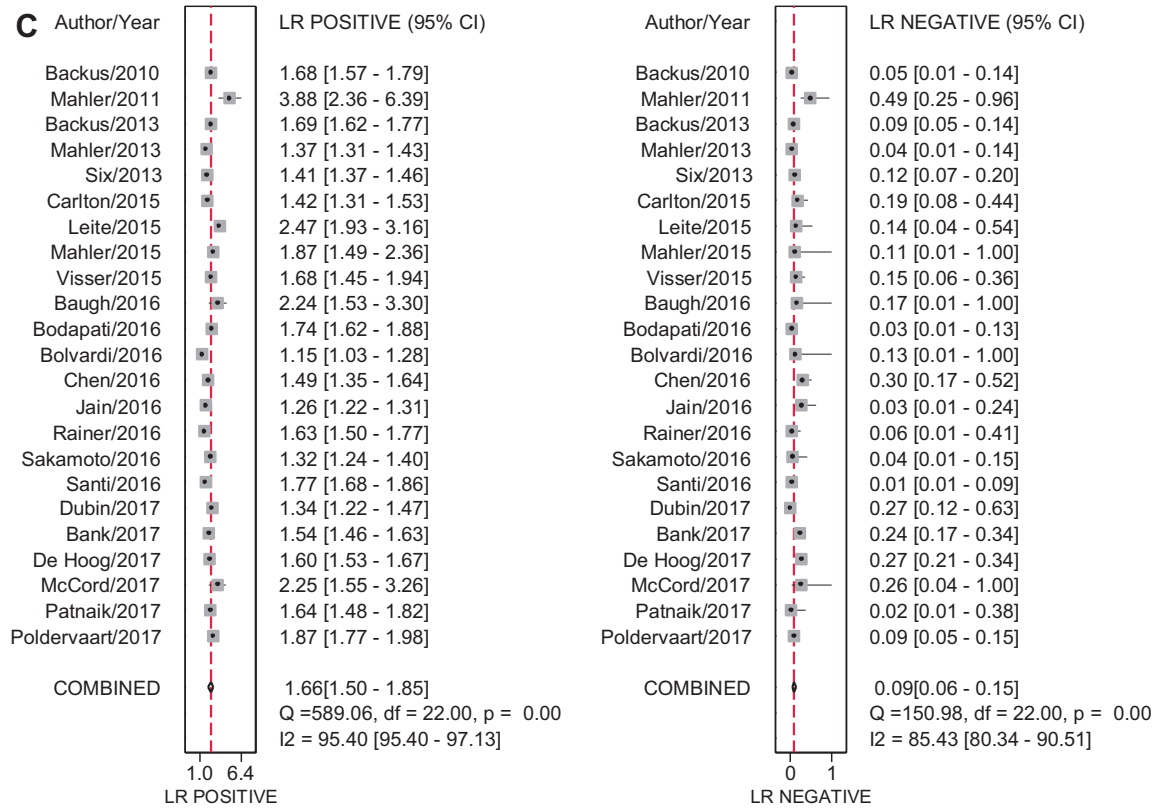

Figure 2. (*continued*).

pooled sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio, are shown in Tables 3 and 4. Briefly, short-term major adverse cardiac event occurrence in North American patients was the lowest (0.7%; 13/1,745) among the cohorts (Table 3). Additionally, using the modified low-risk HEART criteria yielded lower short-term major adverse cardiac event occurrence (0.8%; 21/2,590) (Table 3). However, 67% of studies (6/9) that used modified low-risk HEART criteria occurred in North America (Table 2).

Three distinct geographic regions were identified and included in subgroup analysis: 8 studies enrolled patients from North America, 9 from Europe, and 6 from Asia-Pacific. One study included a mix of patients from both Europe and Asia-Pacific, which we were able to analyze separately.[20] One study included a mix of patients from different regions, which we were unable to analyze separately.[22] Lowest short-term major adverse cardiac event occurrence among patients with low-risk HEART scores was found in the North America group (0.7%; 13/1,745) compared with the European (2.5%; 105/4,230) and Asia-Pacific (2.6%; 63/2,440) patient populations (Table 3). Pooled sensitivities, specificities, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio among different regions are shown in Table 4.

As previously mentioned, a modified low-risk HEART score is defined as one of 0 to 3, including patients with only negative troponin results. Patients with a positive troponin result were considered to have a non–low-risk score, regardless of their initial HEART scores. Of the 25 studies included in the meta-analysis, 9 can be analyzed as using modified low-risk HEART score criteria.[14,21-28] Six of the 9 studies were composed solely of patient populations from North America,[14,23-26,28] 2 were of patients from Europe,[21,27] and 1 was a mixed patient population.[22] Overall, short-term major adverse cardiac event occurrence was lower when a modified low-risk HEART score was used (0.8%; 21/2,590) than when the traditional low-risk HEART score was used (2.4%; 174/7,175) (Table 3). Higher pooled negative predictive value with lower negative likelihood ratio occurred when a modified low-risk HEART score was used (Table 4). However, pooled sensitivity decreased from 97% with the traditional low-risk HEART tool to 93% with the modified one. Given this unexpected result, we performed a post hoc analysis after removal of articles that included only patients with a negative first troponin result because these studies tended to have a much lower sensitivity caused by inclusion

of a lower-risk population. After removal of these 3 studies,[22-24] the sensitivity of the modified low-risk HEART score was 97% (Tables E4 and E5, available online at http://www.annemergmed.com).

Five studies evaluated the HEART score with high-sensitivity troponin reagents.[19,22,48,51,55] Of the 25 studies included in the meta-analysis, 2 used a combination of high-sensitivity troponin and conventional troponin because of inclusion of a multicenter study design within the individual trials.[20,21] Two other studies did not specifically address troponin reagent use.[17,18] A detailed description of troponin study-specific assays is shown in Table 2. The short-term major adverse cardiac event occurrence was lower (0.8%; 12/1,539) in patients with low-risk HEART scores using high-sensitivity troponin compared with ones using conventional troponin reagent (1.8%; 93/5,093) (Table 3). In addition, lower negative likelihood ratio was also shown when use of high-sensitivity troponin for low-risk HEART score calculation occurred (Table 4).

Most studies reported HEART scores relative to short-term (30 days to 6 weeks) major adverse cardiac event outcomes. Two studies reported long-term (≥3 months) major adverse cardiac event outcomes[23,27] and 3 reported both short- and long-term ones.[18,54,55] One study reported 1- and 5-year mortality without addressing composite major adverse cardiac event outcomes,[50] and another reported 1-year major adverse cardiac event outcomes[12] but did not stratify major adverse cardiac events by HEART score. Neither is included in this meta-analysis because of missing data. The major adverse cardiac event occurrence was lower if followed up at a longer interval for patients with low-risk HEART scores (1.1%; 23/2,012) but higher for patients with non–low-risk HEART scores (23.7%; 449/1,894) (Table 3). Pooled sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of low-risk HEART scores to predict short- or long-term major adverse cardiac event outcomes are shown in Table 4.

Agreement on quality assessment was high (first-round screening κ=0.80 [95% CI 0.69 to 0.91]; second-round screening κ=0.82 [95% CI 0.74 to 0.89]). Summaries of the quality assessment of all 31 studies included in the qualitative synthesis are shown in Tables E6 and E7, available online at http://www.annemergmed.com. Studies not qualified by either quality assessment tool were excluded, resulting in a total of 25 studies in the final meta-analysis.

To evaluate publication bias and small-study effects, we applied 3 risk-of-bias tests. First, we performed a Harbord test.[36] The estimated intercept was 1.09, with a standard error of 0.54. Given a $P>.05$, the Harbord test found no statistically significant evidence of publication bias or small-study effects. The corresponding modified Galbraith plot is shown in Table E8, available online at http://www.annemergmed.com. The Peters test had a $P>.05$, which was consistent with our findings from the Harbord test (Table E8, available online at http://www.annemergmed.com).

The Duval and Tweedie trim and fill method yielded a total of 8 imputed studies that were incorporated for analysis. The addition of these 8 missing studies changed the random-effects summary estimate (all-timeframe major adverse cardiac events) from 2.59 (95% CI 2.24 to 2.94) to 2.18 (95% CI 1.83 to 2.53). Although the new estimate was slightly lower, it remained statistically significant ($P<.001$) (Table E8, available online at http://www.annemergmed.com). As a result, the correction for publication bias did not change the overall interpretation of the data; therefore, evidence of publication bias remained low. As expected, the addition of missing studies resulted in an increased variance between studies, subsequently increasing heterogeneity (increased Q statistic from 77 to 117; $P<.001$) (Table E8, available online at http://www.annemergmed.com). The funnel plot shows the imputed data by a square around the data symbol, and the corrected data set looks much more symmetric than the original data set, further indicating no evidence of publication bias (Table E8, available online at http://www.annemergmed.com).

LIMITATIONS

There are several limitations in this study. First, we were not able to include all studies in our meta-analysis, including those published in foreign languages. The exclusion of non-English published studies may have directly affected our subgroup analysis of diverse patient populations across the 3 geographic regions. Second, study eligibility criteria were broad, which may have introduced variations in calculating HEART scores not originally accounted for, and the prevalence of major adverse cardiac events in individual studies varied, indicating different patient populations enrolled in studies and thereby introducing potential variation affecting heterogeneity across studies. High heterogeneity occurred across studies, similar to that in a previously reported meta-analysis.[13] Such high heterogeneity might affect accurate interpretation of results. However, calculating HEART scores with minor deviations from the original derivation[12] has commonly been reported in the literature. Narrow study eligibility criteria would have considerably reduced the number of eligible studies in the final meta-analysis, subsequently introducing greater bias (eg, selection and publication biases). Third, our meta-analysis included

studies with moderate-risk bias, potentially affecting overall results and increasing heterogeneity. Fourth, we did not receive responses from all corresponding authors about requests for raw study data when published study data were missing or unclear. Fifth, most studies did not separate index from nonindex major adverse cardiac events. Subsequently, we did not analyze index major adverse cardiac events in this meta-analysis, which may also have affected the final diagnostic accuracy. Sixth, we were unable to address all potential clinical heterogeneity and spectrum bias for the population included. We reported only diagnostic and prognostic performance among patients with low-risk HEART scores. Therefore, our findings may be skewed and may not adequately summarize overall HEART score performance for risk stratification of ED chest pain patients.

DISCUSSION

Chest pain is one of the most common complaints among ED patients. Failure to appropriately investigate a given patient's chest pain cause presents the risk of sustaining missed acute coronary syndrome, with resulting significantly higher morbidity and mortality.[56] This meta-analysis presents consistent results in regard to low-risk HEART scores predictive of short-term major adverse cardiac events, with a pooled sensitivity of 96% and negative likelihood ratio of 0.09. Our study showed that lower major adverse cardiac event outcomes occurred among both North American patients (0.7%) and those with a modified low-risk HEART score (0.8%). Application of the modified low-risk HEART score has been predominantly studied in North America. Given that the traditional low-risk HEART score has a short-term major adverse cardiac events miss rate of 2.4% and a pooled negative likelihood ratio of 0.08, use of a modified low-risk HEART score can effectively decrease the posttest major adverse cardiac events probability by at least 45% of that expected in a simplified likelihood ratio calculation,[57] which is consistent with our study findings (short-term major adverse cardiac events miss rate of 0.8% when a modified low-risk HEART score was used). Additionally, patients whose low-risk HEART score incorporates use of high-sensitivity troponin are noted to have lower major adverse cardiac event outcomes (0.8%), with a lower pooled negative likelihood ratio (0.07). This trend has been noted among studies performed outside of North America. Therefore, we suggest that chest pain patients stratified as being at low risk according to either a modified low-risk HEART score or a traditional one that incorporates high-sensitivity troponin might be safely discharged directly from the ED without further emergency cardiac testing. Patients should expect a safety window of 30 days to 6 weeks, with minimal risk for progressive morbidity and mortality, to arrange and complete appropriate cardiac follow-up.

Hypothetically, low-risk HEART criteria miss approximately 2% of patients who progress to major adverse cardiac event outcomes. One study reported that emergency physicians might accept a 1% major adverse cardiac events miss rate as a routine practice expectation, with North American physicians being the most tolerant (up to 2%) and Australian physicians the least tolerant (0.1%).[16] Considering this, it appears that overall predicted major adverse cardiac event outcomes exceed physicians' level of acceptance; hence, the traditional HEART score provides less clinical value for ED patients. Because major adverse cardiac events decrease to approximately 1% for patients stratified by either incorporation of high-sensitivity troponin into the traditional HEART calculation or use of the modified low-risk HEART score, ED providers may reach their level of acceptance to discharge such patients.

Our study has findings similar to those previously reported by Van Den Berg and Body,[13] with some differences. We expanded our analysis by adding more contemporary studies because of a surge of recent publications focusing on the use of a low-risk HEART score to risk stratify ED chest pain patients. We enrolled greater than three quarters of screened studies (25/31) into the current meta-analysis, whereas less than 50% (8/25) in the study by Van Den Berg and Body[13] were enrolled. We included studies specifically using modified low-risk HEART criteria because we believe such studies more closely represent current ED practice in North America. Additionally, our results showed that use of modified low-risk HEART criteria yields relatively lower major adverse cardiac event outcomes, falling within the physician level of acceptance expressed in North America. Furthermore, the subgroup analysis of patients whose low-risk HEART scores were calculated with high-sensitivity troponin also yielded lower major adverse cardiac event outcomes. Each of these applications appears to provide an optimal pathway for risk stratification of ED chest pain patients within the current environment. Similar to the previous report, this meta-analysis yielded high statistical heterogeneity among enrolled studies.

Our analysis of different patient populations further validates the generalizability of using the HEART score in diverse patient populations. Moreover, several studies have extended their follow-up intervals, fluctuating from 3 months to 1 year, without affecting sensitivity and negative likelihood ratio, respectively. Increasing follow-up intervals indicate the potential of the HEART score in terms of safely and accurately predicting long-term major adverse cardiac event outcomes.[18,23,27,54,55] Yet, with scant evidence-based studies, we are unable to draw firm conclusions. Because significant heterogeneity occurred

across the studies included in this meta-analysis, future large-scale prospective studies should focus on the use of the HEART score with high-sensitivity troponin to rigorously screen chest pain patients at risk of potential acute coronary syndrome; on post–ED-discharge follow-up interval expansion of short- to long-term major adverse cardiac events; and on cost-savings analysis of health resource allocation.

In this meta-analysis, despite its use in different patient populations, troponin used, and timeline of follow-up, a low-risk HEART score had high sensitivity, negative predictive value, and negative likelihood ratio for predicting major adverse cardiac events, although risk of bias and heterogeneity were high. Use of high-sensitivity troponin in the HEART score, or use of the modified HEART score, may address issues limiting use of the tool in clinical practice.

The authors acknowledge the corresponding authors who accommodated their request for additional study data, especially Timothy Rainer, Edward Carlton, and Leonie Rose Bovino; and Teresa Le, who helped organize data collection and presented the study abstract at the Society for Academic Emergency Medicine annual conference.

Supervising editor: Clare L. Atzema, MD, MSc. Specific detailed information about possible conflict of interest for individual editors is available at https://www.annemergmed.com/editors.

Author affiliations: From the Department of Emergency Medicine, Integrative Emergency Services (Laureano-Phillips, Robinson, Wilson, Schrader, Zenarosa, Wang) and Centers for Outcome Research (Blair), John Peter Smith Health Network (Boyd, Wilson, Schrader, Zenarosa, Robinson, Wang), Fort Worth, TX; the Department of Biostatistics and Epidemiology (Aryal), and the Department of Medical Education, University of North Texas Health Science Center, Fort Worth, TX (Robinson).

Author contributions: JL-P, KB, and HW were responsible for the design search strategy. JL-P and HW were responsible for the inclusion and exclusion criteria and the development of the initial quality and risk-assessment tool. JL-P and KB were responsible for the electronic and manual search. JL-P, DW, and HW were responsible for the study screening, data extraction, and quality and risk-of-bias assessment. JL-P was responsible for contacting study authors for additional data, and for protocol registration. JL-P, RDR, DW, and HW were responsible for writing the article. All authors were responsible for editing the article. SA, SB, and HW were responsible for data analysis. KB was responsible for conduct of the search strategy. HW takes responsibility for the paper as a whole.

All authors attest to meeting the four ICMJE.org authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). The authors have stated that no such relationships exist.

Publication dates: Received for publication February 19, 2018. Revisions received June 27, 2018; September 24, 2018, and November 12, 2018. Accepted for publication December 5, 2018.

Presented as an abstract at the Society for Academic Emergency Medicine annual conference, May 2018, Indianapolis, IN.

REFERENCES

1. Owens PL, Barrett ML, Gibson TB, et al. Emergency department care in the United States: a profile of national data sources. *Ann Emerg Med*. 2010;56:150-165.
2. Goodacre S, Cross E, Arnold J, et al. The health care burden of acute chest pain. *Heart*. 2005;91:229-230.
3. Chase M, Robey JL, Zogby KE, et al. Prospective validation of the Thrombolysis in Myocardial Infarction risk score in the emergency department chest pain population. *Ann Emerg Med*. 2006;48:252-259.
4. Pines JM, Isserman JA, Szyld D, et al. The effect of physician risk tolerance and the presence of an observation unit on decision making for ED patients with chest pain. *Am J Emerg Med*. 2010;28:771-779.
5. Fleischmann KE, Goldman L, Johnson PA, et al. Critical pathways for patients with acute chest pain at low risk. *J Thromb Thrombolysis*. 2002;13:89-96.
6. Lyon R, Morris AC, Caesar D, et al. Chest pain presenting to the emergency department—to stratify risk with GRACE or TIMI? *Resuscitation*. 2007;74:90-93.
7. Cordero A, Rodriguez-Manero M, Garcia-Acuna JM, et al. Additive value of the CRUSADE score to the GRACE score for mortality risk prediction in patients with acute coronary syndromes. *Int J Cardiol*. 2017;245:1-5.
8. Iscanli MD, Metin AN, Evranos B, et al. Comparison of TIMI and Gensini score in patients admitted to the emergency department with chest pain, who underwent coronary angiography. *Med Sci Monit*. 2014;20:343-349.
9. Cullen L, Greenslade J, Hammett CJ, et al. Comparison of three risk stratification rules for predicting patients with acute coronary syndrome presenting to an Australian emergency department. *Heart Lung Circ*. 2013;22:844-851.
10. Boubaker H, Beltaief K, Grissa MH, et al. Inaccuracy of Thrombolysis in Myocardial Infarction and Global Registry in Acute Coronary Events scores in predicting outcome in ED patients with potential ischemic chest pain. *Am J Emerg Med*. 2015;33:1209-1212.
11. Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Heart J*. 2008;16:191-196.
12. Backus BE, Six AJ, Kelder JC, et al. Chest pain in the emergency room: a multicenter validation of the HEART score. *Crit Pathw Cardiol*. 2010;9:164-169.
13. Van Den Berg P, Body R. The HEART score for early rule out of acute coronary syndromes in the emergency department: a systematic review and meta-analysis. *Eur Heart J Acute Cardiovasc Care*. 2018;7:111-119.
14. Mahler SA, Riley RF, Hiestand BC, et al. The HEART Pathway randomized trial: identifying emergency department patients with acute chest pain for early discharge. *Circ Cardiovasc Qual Outcomes*. 2015;8:195-203.

15. Backus BE, Six AJ, Doevendans PA, et al. Prognostic factors in chest pain patients: a quantitative analysis of the HEART score. *Crit Pathw Cardiol*. 2016;15:50-55.
16. Than M, Herbert M, Flaws D, et al. What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the emergency department? a clinical survey. *Int J Cardiol*. 2013;166:752-754.
17. Bolvardi E, Raoufi P, Vak V, et al. Evaluation efficacy of HEART score in prediction of major advanced cardiac events in patients with chest pain. *Biosci Biotech Res Asia*. 2016;13:999-1005.
18. Patnaik S, Shah M, Alhamshari Y, et al. Clinical utility of the HEART score in patients admitted with chest pain to an inner-city hospital in the USA. *Coron Artery Dis*. 2017;28:336-341.
19. Carlton EW, Khattab A, Greaves K. Identifying patients suitable for discharge after a single-presentation high-sensitivity troponin result: a comparison of five established risk scores and two high-sensitivity assays. *Ann Emerg Med*. 2015;66:635-645.
20. de Hoog VC, Lim SH, Bank IE, et al. HEART score performance in Asian and Caucasian patients presenting to the emergency department with suspected acute coronary syndrome. *Eur Heart J Acute Cardiovasc Care*. 2017;7:591-601.
21. Poldervaart JM, Reitsma JB, Backus BE, et al. Effect of using the HEART score in patients with chest pain in the emergency department: a stepped-wedge, cluster randomized trial. *Ann Intern Med*. 2017;166:689-697.
22. McCord J, Cabrera R, Lindahl B, et al. Prognostic utility of a modified HEART score in chest pain patients in the emergency department. *Circ Cardiovasc Qual Outcomes*. 2017;10.
23. Wang H, Watson K, Robinson RD, et al. Chest pain risk scores can reduce emergent cardiac imaging test needs with low major adverse cardiac events occurrence in an emergency department observation unit. *Crit Pathw Cardiol*. 2016;15:145-151.
24. Mahler SA, Hiestand BC, Goff DC Jr, et al. 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;10:128-133.
25. Mahler SA, Miller CD, Hollander JE, et al. Identifying patients for early discharge: performance of decision rules among patients with acute chest pain. *Int J Cardiol*. 2013;168:795-802.
26. Baugh CW, Greenberg JO, Mahler SA, et al. Implementation of a risk stratification and management pathway for acute chest pain in the emergency department. *Crit Pathw Cardiol*. 2016;15:131-137.
27. Melki D, Jernberg T. HEART score: a simple and useful tool that may lower the proportion of chest pain patients who are admitted. *Crit Pathw Cardiol*. 2013;12:127-131.
28. Dubin J, Kiechle E, Wilson M, et al. Mean HEART scores for hospitalized chest pain patients are higher in more experienced providers. *Am J Emerg Med*. 2017;35:122-125.
29. Fesmire FM, Martin EJ, Cao Y, et al. Improving risk stratification in patients with chest pain: the Erlanger HEARTS3 score. *Am J Emerg Med*. 2012;30:1829-1837.
30. Thygesen K, Alpert JS, Jaffe AS, et al. Third universal definition of myocardial infarction. *Eur Heart J*. 2012;33:2551-2567.
31. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158:280-286.
32. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11:e1001744.
33. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924-926.
34. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529-536.
35. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ*. 2006;333:597-600.
36. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*. 2006;25:3443-3457.
37. Peters JL, Sutton AJ, Jones DR, et al. Comparison of two methods to detect publication bias in meta-analysis. *JAMA*. 2006;295:676-680.
38. Egger M, Davey SG, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-634.
39. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med*. 1988;7:889-894.
40. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56:455-463.
41. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001;20:2865-2884.
42. Gagnier JJ, Morgenstern H, Altman DG, et al. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. *BMC Med Res Methodol*. 2013;13:106.
43. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264-269, W64.
44. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) Group. *JAMA*. 2000;283:2008-2012.
45. Backus BE, Six AJ, Kelder JC, et al. A prospective validation of the HEART score for chest pain patients at the emergency department. *Int J Cardiol*. 2013;168:2153-2158.
46. Six AJ, Cullen L, Backus BE, et al. The HEART score for the assessment of patients with chest pain in the emergency department: a multinational validation study. *Crit Pathw Cardiol*. 2013;12:121-126.
47. Leite L, Baptista R, Leitao J, et al. Chest pain in the emergency department: risk stratification with Manchester triage system and HEART score. *BMC Cardiovasc Disord*. 2015;15:48.
48. Visser A, Wolthuis A, Breedveld R, et al. HEART score and clinical gestalt have similar diagnostic accuracy for diagnosing ACS in an unselected population of patients with chest pain presenting in the ED. *Emerg Med J*. 2015;32:595-600.
49. Bodapati SN, Gunnarsson R, McBride WJ, et al. Chest pain risk assessment in Indigenous and non-Indigenous Australians using HEART score. *Emerg Med Australas*. 2016;28:138-144.
50. Jain T, Nowak R, Hudson M, et al. Short- and long-term prognostic utility of the HEART score in patients evaluated in the emergency department for possible acute coronary syndrome. *Crit Pathw Cardiol*. 2016;15:40-45.
51. Rainer TH, Leung YK, Lee A, et al. Add-on tests for improving risk-stratification in emergency department patients with chest pain who are at low to moderate risk of 30-day major adverse cardiac events. *Int J Cardiol*. 2016;220:299-306.
52. Sakamoto JT, Liu N, Koh ZX, et al. Comparing HEART, TIMI, and GRACE scores for prediction of 30-day major adverse cardiac events in high acuity chest pain patients in the emergency department. *Int J Cardiol*. 2016;221:759-764.
53. Bank IEM, de Hoog VC, de Kleijn DPV, et al. Sex-based differences in the performance of the HEART score in patients presenting to the emergency department with acute chest pain. *J Am Heart Assoc*. 2017;6.
54. Chen XH, Jiang HL, Li YM, et al. Prognostic values of 4 risk scores in Chinese patients with chest pain: prospective 2-centre cohort study. *Medicine (Baltimore)*. 2016;95:e4778.
55. Santi L, Farina G, Gramenzi A, et al. The HEART score with high-sensitive troponin T at presentation: ruling out patients with chest pain in the emergency room. *Intern Emerg Med*. 2017;12:357-364.
56. Fox KA, Carruthers KF, Dunbar DR, et al. Underestimated and under-recognized: the late consequences of acute coronary syndrome (GRACE UK-Belgian Study). *Eur Heart J*. 2010;31:2755-2764.
57. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17:646-649.