Original Contribution

Comparison of clinical risk scores for triaging high-risk chest pain patients at the emergency department

Salah S. Al-Zaiti, PhD a,b,c, Ziad Faramand, MD a,e, Mohammad O. Alrawashdeh, PhD a,f, Susan M. Sereika, PhD c,d, Christian Martin-Gill, MD, MPH b,c,e, Clifton Callaway, MD, PhD b,e

a Department of Acute & Tertiary Care Nursing, University of Pittsburgh, PA, United States
b Department of Emergency Medicine, University of Pittsburgh, PA, United States
c Department of Research & Evaluation, University of Pittsburgh, PA, United States
d Department of Biostatistics, University of Pittsburgh, PA, United States
e University of Pittsburgh Medical Center (UPMC), Pittsburgh, PA, United States
f University of Pittsburgh School of Nursing, 3500 Victoria Street, Pittsburgh, PA 15261, United States.

Abstract

Background: Many of the clinical risk scores routinely used for chest pain assessment have not been validated in patients at high risk for acute coronary syndrome (ACS). We performed an independent comparison of HEART, TIMI, GRACE, FRISC, and PURSUIT scores for identifying chest pain due to ACS and for predicting 30-day death or re-infarction in patients arriving through Emergency Medical Services (EMS).

Methods and results: We enrolled consecutive EMS patients evaluated for chest pain at three emergency departments. A reviewer blinded to outcome data retrospectively reviewed patient charts to compute each risk score. The primary outcome was ACS diagnosed during the primary admission, and the secondary outcome was death or re-infarction within 30-days of initial presentation. Our sample included 750 patients (aged 59 ± 17 years, 42% female), of whom 115 (15.3%) had ACS and 33 (4.4%) had 30-day death or re-infarction. The c-statistics of HEART, TIMI, GRACE, FRISC, and PURSUIT for identifying ACS were 0.87, 0.86, 0.73, 0.84, and 0.79, respectively, and for predicting 30-day death or re-infarction were 0.70, 0.73, 0.72, 0.72, and 0.62, respectively. Sensitivity/negative predictive value of HEART ≥ 4 and TIMI ≥ 3 for ACS detection were 0.94/0.98 and 0.87/0.97, respectively. Conclusions: In chest pain patients admitted through EMS, HEART and TIMI outperform other scores for identifying chest pain due to ACS. Although both have similar negative predictive value, HEART has better sensitivity and lower rate of false negative results, thus it can be used preferentially over TIMI in the initial triage of this population.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Nearly 7 million Americans visit emergency departments (ED) annually with a chief complaint of chest pain. The focus of initial chest pain assessment is to identify those with acute coronary syndrome (ACS) to accelerate necessary treatment. Given that the initial clinical examination, ECG results, and initial cardiac enzyme levels are often insufficient to eliminate the possibility of ACS, a large proportion of chest pain patients are admitted to the hospital for further evaluation [1–3]. Many well-established clinical risk scores (e.g., HEART, TIMI, etc.) have been used to stratify risk in such undifferentiated chest pain populations [4], and the HEART score has been shown to possess the highest negative predictive value for adverse cardiac events compared to other risk scores [5]. However, nearly one quarter of chest pain patients are admitted through Emergency Medical Services (EMS). This latter group tends to have higher prevalence of ACS events and subsequent hospital readmission, as well as three-fold increased risk of death compared to those who self-transport [6–8]. As such, clinicians have very low threshold to admitting these patients for further evaluation. Despite the established value of clinical risk scores in stratifying risk in unselected ED populations, these risk scores were not specifically validated in high risk chest pain patients admitted through EMS. Still, these risk scores are frequently associated with prognosis at 30 days.

Abbreviations: ACS, acute coronary syndrome; ECG, electrocardiogram; EMS, Emergency Medical Services; ED, emergency department; FRISC, fast revascularisation in instability in coronary disease; GRACE, global registry of acute coronary events; HEART, history-ECG-age-risk factors-troponin; MACE, major adverse cardiac events; PURSUIT, platelet glycoprotein IIb/IIIa in unstable angina: receptor suppression using integrilin therapy; TIMI, thrombolysis in myocardial infarction.

Corresponding author at: Department of Acute and Tertiary Care, University of Pittsburgh, School of Nursing, 3500 Victoria Street, Pittsburgh, PA 15261, United States.
E-mail address: ss333@pitt.edu (S.S. Al-Zaiti).

https://doi.org/10.1016/j.ajem.2018.06.020
0735-6757/© 2018 Elsevier Inc. All rights reserved.
rather than with diagnostic performance during hospitalization. Accordingly, we sought to perform a comparison of PURSUIT, TIMI, GRACE, FRISC, and HEART risk scores for (1) identifying chest pain due to ACS, (2) predicting the occurrence of subsequent cardiac events, and (3) predicting all-cause readmissions.

2. Methods

2.1. Sample and settings

Subjects for this sub-analysis were recruited from the EMPiRE study (ECG Methods for the Prompt Identification of Coronary Events). The methods of EMPiRE were described in detail elsewhere [9]. In short, EMPiRE is a prospective observational cohort study that enrolled consecutive, non-traumatic chest pain patients transported by EMS to one of three University of Pittsburgh Medical Center-affiliated tertiary care hospitals (UPMC Presbyterian, Mercy, and Shadyside hospitals) in Pittsburgh, PA between May 2013 and August 2014 (N = 2065). The current secondary analysis is based on patients on whom a 12-lead ECG was obtained upon first medical contact by EMS and transmitted to the UPMC medical command center as part of the acute evaluation of chest pain (n = 750). Independent reviewers abstracted the key in-hospital data elements recommended by the American College of Cardiology for measuring the management and outcomes of patients with ACS [10]: demographics, anthropometrics, past medical history, home medications, clinical presentation and course of hospitalization, laboratory tests, imaging studies, cardiac catheterization, treatments, and in-hospital complications. There were no modifications to routine medical care. Patients were enrolled through a waiver of informed consent and the study was approved by the University of Pittsburgh Institutional Review Board.

2.2. Clinical outcomes

The first primary outcome was the presence of ACS (myocardial infarction or unstable angina) during the primary indexed admission, defined as the presence of symptoms of ischemia (i.e., diffuse discomfort in the chest, upper extremity, jaw, or epigastric area for >20 min) and at least one of the following criteria: (1) elevation of cardiac troponin (i.e., >99th percentile), (2) subsequent development of lable, ischemic ECG changes (e.g., ST-T changes, new bundle branch block, new Q wave) during hospitalization, (3) imaging evidence of new loss of viable myocardium or new regional wall motion abnormalities, or (4) coronary angiography or nuclear imaging demonstrating >70% stenosis of a major coronary artery with or without treatment [11,12]. The second primary outcome was the occurrence of major adverse cardiac events (MACE) within 30 days of presentation, defined as (1) resuscitated or unresuscitated sudden cardiac arrest, (2) all-cause death, or (3) post-discharge re-infarction requiring cardiac revascularization. The secondary outcome was all-cause readmission or observation stays within 30 days of initial presentation.

To adjudicate study outcomes, two independent reviewers were provided full access to patient index and discharge records, serial ECGs, results of cardiac diagnostic tests (e.g., imaging scans and catheterization laboratory reports), and other information pertinent to the course of hospitalization (e.g., interventions, procedures, and prescribed medications). Patients who did not have complete evaluation within 30 days (e.g., discharged from ED) were classified as negative for MACE and ACS if they had no 30-day adverse events. Patients who died without explanation within 30 days were classified as positive for both outcomes. All disagreements were resolved by a third reviewer.

2.3. Calculating risk scores

Risk scores were calculated retrospectively through chart review by an independent reviewer who was blinded to outcome data. We followed the methods described in the derivation studies for calculating each risk score [4,5,13-17]. Each of the five risk scores used some criteria based on (1) cardiac enzymes and (2) the presenting 12-lead ECG. We used the first cardiac troponin I result at the ED, with the 99th percentile cutoff point specified at each participating UPMC hospital. As recommended by current guidelines for the management of patients with suspected ACS, we used the initial prehospital 12-lead ECG upon first medical contact. The ECG item in each risk score was given a full score if it met the diagnostic cutoff values specified in the universal definition of myocardial infarction: [12] two contiguous leads with (1) ST elevation ≥0.2 mV (in V2–V3) or ≥0.1 mV (in other leads), (2) horizontal/down-sloping ST depression ≥0.5 mV, or (3) T-wave inversion ≥0.1 mV. If the ECG revealed a bundle branch block, a paced rhythm, a ventricular rhythm or left ventricular hypertrophy (based on Sokolow-Lyon voltage criterion), then the ECG item was not given a score unless it was specified in that risk score (i.e., HEART gives partial ECG score). All other scoring variables (i.e., age, sex, past medical history, clinical presentation, and diagnostic lab results) were obtained from the in-hospital records. A brief description of each risk score and with any adjustments we made is summarized in Table 1.

2.4. Statistical analysis

For continuous type variables, measures of central tendency and dispersion were reported as mean ± SD for normally distributed variables and median [25th–75th IQR] for non-normally distributed variables. Categorical variables were described using frequencies and percentages and reported as n (%). Each outcome was treated as a dichotomous variable (yes/no) and each risk score was treated as a continuous variable in analyses. ACS groups were compared using the chi-square test of independence for categorical variables, the independent samples t-test for normally distributed continuous variables, and the Mann-Whitney test for non-normally distributed continuous variables. The discriminatory power to identify each primary outcome using each risk score was evaluated using the area under the receiver-operator characteristic curve. Areas under the ROC curve (AUC) for the different risk scores were compared using the nonparametric method developed by Delong, DeLong, and Clarke-Pearson with post-hoc multiple pairwise comparisons at Bonferroni adjustment at p < 0.005 (=0.05 / 10) [18]. Previously published cutoff limits for each risk score were used to compute and report diagnostic accuracy values as per STARD guidelines [19], including sensitivity, specificity, as well as positive (PPV) and negative (NPV) predictive values. We used 1000 samples bootstrap technique for generating the 95% CI limits for each diagnostic accuracy value. The pairwise

Table 1

<table>
<thead>
<tr>
<th>Scoring criteria for computing each clinical risk score.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Past medical history</td>
</tr>
<tr>
<td>Risk factors</td>
</tr>
<tr>
<td>Medication use</td>
</tr>
<tr>
<td>CSS/Killip class</td>
</tr>
<tr>
<td>Signs &amp; symptoms</td>
</tr>
<tr>
<td>Patient history</td>
</tr>
<tr>
<td>Cardiac arrest upon admission</td>
</tr>
<tr>
<td>Vital signs</td>
</tr>
<tr>
<td>ECG findings</td>
</tr>
<tr>
<td>Cardiac enzymes</td>
</tr>
<tr>
<td>Creatinine level</td>
</tr>
<tr>
<td>Possible range of scores</td>
</tr>
<tr>
<td>Cutoff for high-risk</td>
</tr>
</tbody>
</table>

A FRISC score uses interleukin-6 or CRP in the scoring criteria, which were missing from our data, yielding a possible score of 0–5 rather than 0–6 as described in the original method.
agreement between different scores was evaluated using the McNemar's test. The significance level was set at 0.05 for two-tailed hypothesis testing. All analyses were conducted using IBM SPSS Statistics version 24 (IBM Corp., Armonk, NY) and SAS version 9.4 (SAS Institute, Inc., Cary, NC).

3. Results

The final sample included 750 patients (age 59 ± 17 years; 42% female, and 40% Black). Hypertension (68%) and smoking (58%) were the most prevalent risk factors for ACS, followed by dyslipidemia (35%), known CAD (33%), prior MI (27%), prior revascularization procedure (28%), diabetes (26%), and heart failure (18%). Table 2 summarizes the demographic and clinical characteristics of the study sample.

Overall, 115 patients (15.3%) were diagnosed with ACS during the primary admission (31 had STE on presenting ECG and 84 had NSTE). Compared to those with no ACS, patients with ACS were older and were more likely to have a prior history of diabetes mellitus, prior MI, or prior revascularization procedure (Table 2). The initial cardiac troponin was positive in only 50% of patients with ACS; repeated troponin essays were positive in 87% of this group (i.e., 13% had unstable angina). More than half (53%) of ACS cases were treated by primary percutaneous coronary intervention, and only 6% were treated by coronary bypass surgery.

A total of 42 MACE events occurred in 33 patients (4.4%) within 30-days: 12 developed fatal ventricular arrhythmia upon admission; 4 experienced post-discharge new infarctions with unplanned repeat revascularization. There were no differences between patients with or without MACE regarding age, sex, race, or past medical history. The rate of 30-day MACE was 16.5% in ACS patients, compared to 2.2% in others (p < 0.001) (Table 2).

A total of 146 patients (20%) were readmitted to the hospital within 30-days of initial presentation. Approximately, one third (n = 47, 32%) of these readmissions were chest-pain related. Overall, patients who were readmitted were more likely to be males and to have a past medical history of hypertension. There was no difference in 30-day readmission rates between those diagnosed with or without ACS in the primary admission (Table 2).

Fig. 1 shows the differences in risk scores between patients with or without the primary clinical outcomes. TIMI, GRACE, FRISC, and HEART scores were higher in patients with ACS (all p < 0.001), in patients with 30-day MACE (all p < 0.001), and in patients with 30-day readmissions (all p < 0.05). PURSUIT score was higher only in patients with chest pain due to ACS (p < 0.001), but not in patients with 30-day MACE or 30-day readmissions.

Fig. 2 compares the area under the ROC curve between the five risk score. HEART (0.87), TIMI (0.86), and FRISC (0.84) had good discrimination power to classify ACS events; PURSUIT (0.79) and GRACE (0.73) had a fair discrimination power. Overall these areas under the ROC curves differed across the five risk scores (χ² = 38.24, df = 4, p < 0.0001). For classifying 30-day MACE events, however, none of the scores had a good discrimination power. TIMI (0.73), FRISC (0.72), HEART (0.70), and GRACE (0.70) had fair discrimination power; whereas PURSUIT (0.62) had poor discrimination power. These areas under the ROC curves did not significantly differ across the five risk scores (χ² = 5.46, df = 4, p = 0.2443). Using previously published cut-off values, we computed the diagnostic accuracy values for each risk score to classify ACS and 30-day MACE (Table 3). For ACS detection, HEART had the highest sensitivity (0.94) and NPV (0.98); whereas TIMI had lower sensitivity (0.87) but comparable NPV (0.97). Similar results were observed for MACE prediction: HEART had better sensitivity than TIMI (0.79 vs. 0.67), but both had comparable NPV (0.98 vs. 0.98).

Fig. 3 compares the numbers of misclassified events by each risk score. HEART misclassified 7 out of the 115 ACS events (6%) and 7 out of the 33 MACE events (21%); whereas TIMI misclassified 15 (13%) and 11 (33%) of these events, respectively. Also, there was a significant disagreement between how HEART and TIMI scores assigned risk categories (McNemar's χ² = 237.4, df = 1, p < 0.001). Specifically, HEART classified 429 (57%) patients as higher risk, compared to only 269 (36%) patients by TIMI. More than 40% of those classified as higher risk by HEART (i.e., 175 of 429) were classified as lower risk by TIMI. Finally, to evaluate the discrimination index of each risk score to classify NSTE-ACS events, we repeated all analyses after excluding patients with diagnostic ST elevation on the presenting ECG who were directly referred for emergent catheterization (n = 31). We observed similar results; only HEART, TIMI, and FRISC had good discrimination power (Fig. 4). HEART had the highest sensitivity (93%), but the lowest specificity (49%), whereas TIMI had a better balance between sensitivity (85%) and specificity (73%).

4. Discussion

In this study, we compared the performance of five clinical risk scores in high-risk chest pain patients transported by ambulance with an ECG transmitted by EMS to the receiving hospital. We found that, in this EMS population, clinical risk scores perform better in identifying chest pain due to ACS, rather than predicting 30-day MACE. HEART (AUC = 0.87) and TIMI (AUC = 0.86) scores had comparable power and NPV to identify ACS events, but HEART had better sensitivity and lower rate of false negative results. To our knowledge, this is the first study to validate the use of clinical risk scores in a large EMS cohort of patients with chest pain. The originality of our work lies in the comparison of the diagnostic performances of five risk scores for optimal early triage of patients with suspected ACS.
Few studies previously compared the prognostic value of different risk scores for predicting MACE in unselected chest pain populations. These studies consistently favored HEART score over other clinical risk scores. For instance, Poldervaart and colleagues recently compared three risk scores for predicting 6-week MACE (19%) in nine Dutch hospitals (n = 1748) [20]. They found that HEART had the highest discrimination index (AUC = 0.86), followed by TIMI (AUC = 0.80) and GRACE (AUC = 0.73). Sakamoto and colleagues also recently compared the same three risk scores for predicting 30-day MACE (36%) in a single hospital in Singapore (n = 604) [21]. They also found that HEART had the highest discrimination index (AUC = 0.78), followed by TIMI (AUC = 0.65) and GRACE (AUC = 0.62). In a third prospective validation study by the authors of the HEART score [5], the HEART score (AUC = 0.83) outperformed TIMI (AUC = 0.75) and GRACE (AUC = 0.70) in

Fig. 1. Comparing mean risk scores between those with or without major clinical outcomes. This figure compares the mean and its 95% confidence interval for each risk score between patients with or without acute coronary syndrome (ACS), with or without 30-day major adverse cardiac events (MACE), and with or without 30-day readmission.
predicting 6-week MACE (17%) in a large cohort of patients (n = 2440). In a fourth study, Sun and colleagues compared TIMI and HEART risk scores for predicting MACE (6.2%) in nine U.S. hospitals (n = 8255) [22]. They similarly found that HEART score had a higher discrimination index (AUC = 0.75) compared to TIMI (AUC = 0.67). Our current findings are congruent with these previous studies and demonstrate that HEART and TIMI maintain superior discrimination power for identifying ACS events in high-risk EMS cohorts of patients with chest pain.

One important comparison can be made between our data and previous studies. That is we observed relatively better performance for both HEART (0.87) and TIMI (0.86) scores for predicting ACS compared to previous studies (0.75–0.86 for HEART and 0.65–0.80 for TIMI). We speculate that this might be due to the fact that our primary outcome included only patients with ACS diagnosed at the index-visit, whereas previous studies used a composite endpoint of 6-week ACS, invasive coronary revascularization, post-discharge re-infarction, and all-cause death. In fact, when we considered 30-day death or re-infarction the c-statistic dropped significantly from 0.87 to 0.70 for HEART and from 0.86 to 0.73 for TIMI. In other words, the lower c-statistic observed in the previous studies might be attributed to dilution effect introduced by using a soft composite endpoints [23]. Interestingly, one of these previous studies [5] considered the occurrence of ACS within three months (i.e., 93% of events were already diagnosed during the primary admission) as a secondary outcome and found that the c-statistics for HEART score increased from 0.83 for MACE prediction to 0.86 for ACS prediction. Another interesting comparison can be seen in one of these previous studies that excluded patients with ACS diagnosed at the ED from their MACE endpoint [22]. As expected, the c-statistic they observed for HEART (0.75) was relatively lower than what we observed (0.87) and what most other studies previously reported (0.75–0.86).

There are several important clinical implications of our findings. The focus of chest pain assessment in the ED is to identify both low-risk and high-risk patients. HEART and TIMI scores can be computed as soon as the first lab results and initial ECG are obtained (<1 h), and hence can play a significant role in diagnosing or excluding the possibility of ACS. In high-risk patients transported by EMS providers, HEART score is very useful in identifying which patients are more likely to have ACS (warranting admission or further evaluation). Within this context, HEART score also has a very high NPV (i.e., >97%) which suggests it can be used to identify low risk patients who are less likely to experience adverse events. It is worth noting that HEART score tends to classify more patients as high-risk (57%) compared to other scores (36% using TIMI), yielding low specificity (i.e., <50%) and raising concerns regarding increased resource utilization. Nevertheless, the gains associated with lower rates of false negatives would keep this score more

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Diagnostic accuracy of each risk score in predicting study outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
</tr>
<tr>
<td>Predicting ACS</td>
<td></td>
</tr>
<tr>
<td>GRACE</td>
<td>0.73</td>
</tr>
<tr>
<td>PURSUIT</td>
<td>0.79</td>
</tr>
<tr>
<td>FRISC</td>
<td>0.84</td>
</tr>
<tr>
<td>TIMI</td>
<td>0.86</td>
</tr>
<tr>
<td>HEART</td>
<td>0.87</td>
</tr>
<tr>
<td>Predicting MACE</td>
<td></td>
</tr>
<tr>
<td>GRACE</td>
<td>0.72</td>
</tr>
<tr>
<td>PURSUIT</td>
<td>0.62</td>
</tr>
<tr>
<td>FRISC</td>
<td>0.72</td>
</tr>
<tr>
<td>TIMI</td>
<td>0.73</td>
</tr>
<tr>
<td>HEART</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Values are mean ± 95% CI using 1000 samples bootstrap technique; PPV: positive predictive value; NPV: negative predictive value.
† p < 0.005 against HEART.
‡ p < 0.005 against TIMI.
appealing for the development of clinical management algorithms compared to other scores. Still, supplementing HEART score with recent novel tools designed for non-invasive ischemia detection [24-26] would potentially improve the specificity of this score for optimal clinical utility. Finally, given that HEART computation relies more on subjective data, TIMI would provide more objective assessment when clinicians cannot reliably obtain a detailed description of the chief complaint (e.g., patient is poor historian).

4.1. Limitations

One limitation of this analysis is that all risk scores were computed retrospectively, which is especially important for scoring the history component of the HEART score. To address this issue, we used a systematic coding scheme by a reviewer blinded to clinical outcomes, and reviewed all available ED records for complete assessment (i.e., EMS charts, triage reports, ED progress notes, ED nursing notes, admission notes, and ED discharge summary). Another limitation is the possibility of loss to follow up due to the observational nature of our study. We used inpatient and outpatient electronic medical records of UPMC network to assure complete ascertainment, and our 30-day death or re-infarction rate was 4.4%, which is comparable to the rate reported in chest pain populations in the United States using Medicare billing data (i.e., 3.2–4.3%) [27].

5. Conclusion

In high-risk chest pain patients admitted through EMS, clinical risk scores have superior discrimination index for identifying ACS as the cause of chest pain during the primary admission, but poor discrimination index for predicting future death or re-infarction. Although HEART and TIMI have comparable excellent discrimination power, HEART score has better sensitivity and lower rate of false negative results, thus it can be used preferentially over TIMI in the initial ED evaluation of this population.

Fig. 3. Comparing high and low risk categories in patients with major clinical outcomes. This figure compares the number of patients correctly classified as high risk or misclassified as low risk by each risk score for detecting: (A) acute coronary syndrome (ACS), or (B) 30-day major adverse cardiac events (MACE). Error bars indicate the 95% CI of the count.

Fig. 4. Predicting ACS in the absence of diagnostic ST elevation on the presenting ECG. (A) This figure compares the area under the receiver operator characteristics curve (AUC) for each risk score for detecting non-ST elevation acute coronary syndrome (NSTE-ACS). (B) This figure compares the number of patients correctly classified as high risk or misclassified as low risk by HEART and TIMI for detecting NSTE-ACS.
Funding

NIH – NHLBI (R01-HL-137761).

Conflicts of interest

None.

Presentation

This work was presented at the 2017 AHA Scientific Sessions in Anaheim CA.

References


