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

Anniek Visser, Albert Wolthuis, Rob Breedveld, Ewoud ter Avest

What is already known on this subject

The HEART score is a simple bedside scoring system that quantifies risk of ischaemic events in patients with chest pain presenting to the emergency department. In a prospective validation by Backus in 2013, the HEART score was more accurate than the Thrombolysis in Myocardial Infarction and Global Registry in Acute Coronary Events scores in predicting outcome.

What this study adds

In this prospective cohort study of an unselected population of patients with chest pain presenting to an emergency department in the Netherlands, the HEART score and clinical gestalt demonstrated similar diagnostic accuracy for acute coronary syndrome.

ABSTRACT

Background

Acute coronary syndrome (ACS) can be a diagnostic challenge in the emergency department (ED). Recently, the HEART score was developed, a simple bedside scoring system that quantifies risk of ischaemic events in patients with undifferentiated chest pain presenting in the ED.

Objective

In this prospective cohort study, we compared the diagnostic accuracy of HEART score and clinical gestalt (clinical judgement) for diagnosing ACS in an unselected population of patients with chest pain presenting to the ED.

Methods

HEART score (0–10) and clinical gestalt (low risk, intermediate risk or high risk of ACS) were prospectively determined in the ED in 255 patients presenting with chest pain by the treating physician. The reference standard was the presence of ACS, which was defined as either acute myocardial infarction (AMI) or the occurrence of a major adverse cardiac event within 6 weeks after presentation in the ED.

Results

75 out of 255 patients (29%) had an ACS. A HEART score ≤3 had a lower negative likelihood ratio (0.15 (0.06–0.36)) for ACS than a low risk based on clinical gestalt (0.35 (0.19–0.64)), whereas a high HEART score ≥7 had a higher positive likelihood ratio (5.2 (3.2–8.5) vs 3.1 (2.2–4.4)). However, c-statistic of HEART score was not significantly different from clinical gestalt (0.81 (0.76–0.86) vs 0.79 (0.73–0.84), p=0.13).

Conclusions

Our study demonstrates that 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.

INTRODUCTION

Acute coronary syndrome (ACS) can be a true diagnostic challenge. It represents a wide clinical spectrum, ranging from unstable angina to ST elevation myocardial infarction (MI) and, at each stage, there is significant heterogeneity in presentation among individuals. Furthermore, many conditions mimic ACS. Of all patients who present to the emergency department (ED) with chest pain, only 25% has an ACS. A quick, accurate diagnosis, however, is of utmost importance: discharge from the ED in case of an ACS may result in a life-threatening outcome, whereas unnecessary admission of patients results in overuse of ED facilities and hospital facilities and high costs.

Traditionally, ACS diagnosis in the ED is based on three pillars: history, ECG and cardiac biomarker measurements. Numerous risk prediction scores have been developed that incorporate these characteristics to discriminate those patients at high risk for a major adverse cardiac event (MACE) from those with a low risk. The two most commonly used scores are the Global Registry in Acute Coronary Events (GRACE) and the Thrombolysis in Myocardial Infarction (TIMI). However, both were developed for risk stratification of patients with proven ACS, whereas the challenge in the ED is to diagnose ACS. More recently, the HEART score was developed, which is a simple bedside scoring system closely following clinical reasoning. Unlike TIMI and Global Registry in Acute Coronary Events (both developed for prognostication purposes), HEART score aims to improve the accuracy of diagnosing ACS by physicians seeing patients with undifferentiated chest pain in the ED. The HEART score contains five items (history, ECG, age, risk factors and troponin (table 1)). By appreciating each of these five elements with 0, 1 or 2 points, each patient will receive a score of 0–10. The primary end point of the HEART score is a composite of: acute myocardial infarction (AMI), percutaneous coronary intervention (PCI), coronary artery bypass graft (CABG) surgery and death, all occurring within 6 weeks, together called MACE. These outcomes are typically related to an ACS and therefore considered indirect proof of the diagnosis.
However, HEART score incorporates only a selection of the information available to the physician at the moment of presentation of the patient in the ED. Many physicians rely on their clinical judgement (clinical gestalt) when they see patients with chest pain. So far, no study has compared the diagnostic accuracy of HEART score and clinical gestalt in patients with undifferentiated chest pain. Therefore, in this study, we set out to compare the diagnostic accuracy of HEART score and clinical gestalt for diagnosing ACS in an unselected population of patients with chest pain presenting to the ED.

### METHOD

#### Study setting and study population

A single-centre prospective cohort study of all adult patients presenting with chest pain to the ED was performed in the Medical Center Leeuwarden, a teaching hospital with 23,000 ED visits yearly. Data were collected during an 8-month study period (from 1 December 2012 until 31 July 2013). Patients admitted to the ED during the study period with chest pain were eligible for inclusion in the study when they were at least 18 years old and had experienced at least 5 min of chest pain that could be related to an ACS. Patients presenting with only syncope, shortness of breath, dyspnoea, palpitations or atypical complaints like fatigue, nausea or dizziness were excluded, as were patients with an ST elevation MI on the ECG, interhospital referrals and patients that had been included on prior visits. As our study only involved evaluation of routinely recorded patient data, this type of study was determined to be exempt research by our local institutional review board.

#### Data acquisition

During the study, data were collected on a Case Report Form (CRF) by the treating physician while the patient was in the ED. Treating physicians were either emergency medicine consultants or residents (emergency medicine or internal medicine in their 1st year until their 4th year of training). The CRF consisted of separate entries for age, gender, level of training of the physician seeing the patients, medication administration in the ED and final discharge diagnosis from the ED. Furthermore, the treating physician was asked to rate the clinical gestalt based probability (low, intermediate or high) of the presence of an ACS as the cause of the patient’s chest pain symptoms at presentation. All available information (history, physical exam, ECG and laboratory results) could be used by the physician to determine clinical gestalt-based probability. Finally, the HEART score (table 1) and its five component subscores (history, ECG, age, number of risk factors and troponin) had to be filled in. Two points were assigned to the ECG subscore when ST-depression (defined as ST-segment depression ≥0.05 mV in two or more contiguous leads in the appropriate clinical context) was present on the ECG. One point was assigned when non-specific repolarisation disturbances or a left bundle branch block were present. The HEART score and the probability of ACS according to clinical gestalt were assessed as soon as the first lab results and ECG were obtained.

ECGs were recorded for all patients and electronically stored for review purposes. High sensitivity troponin T (hsTropT) measurements were performed in all patients at presentation, with results being available within 60 min (Roche Modular E170, Roche Diagnostics Mannheim, Germany; 99th centile 14 ng/L). When the patient had chest pain for more than 6 h at presentation with a normal ECG and a normal initial hsTropT, no serial hsTropT was ordered, which is in line with the most recent European Society of Cardiology (ESC) guideline recommendations. Otherwise, serial hsTropT was ordered 6 h after presentation. In case of serial troponin measurements, only the troponin T value of the first blood sample was used for the HEART score calculation.

#### Follow-up

Follow-up data were retrieved from the electronic patient records and included discharge letters, revascularisation reports and documentation from follow-up visits in the outpatient clinic.

#### Diagnostic reference standard

The diagnostic reference standard in our study was ACS, which was defined as acute MI at presentation or MACE within 6 weeks of presentation in the ED, since the occurrence of MACE in a patient without MI at presentation would suggest the diagnosis at presentation was ACS with MI. MACE was defined as either MI, or PCI, or CABG, or coronary angiography revealing significant stenosis or death by any cause. AMI was defined according to the ESC-American Heart Association-WHO task force for the redefinition of MI as: ‘a rise and fall of troponin values with at least one value above the 99th centile of the upper reference limit together with evidence of myocardial ischaemia’.

#### Statistical analysis

Descriptive statistics are represented as average (95% CI) or absolute numbers. Differences in prevalence of ACS between various risk groups were tested by one-way analysis of variance with post hoc multiple comparisons (Tukey Kramer). In order to compare diagnostic accuracy of HEART score and clinical gestalt c-statistics were calculated and compared with the Hanley and McNeil method. For all tests, statistical significance was accepted at the 95% CI. All statistical analyses were done with the SPSS V22.0 for Windows statistical package (SPSS, Chicago Illinois, USA) and the MedCalc V12.7.8 statistical program.

### RESULTS

#### Study population and patient characteristics

During the study period, 474 patients visited the ED with chest pain meeting our inclusion criteria. CRF was completed in 279...
of these patients. In 24 patients follow-up data 6 weeks after their visit were unavailable and our study population consisted of the remaining 255 patients. Patient characteristics for this study population are represented in table 2. Mean (95% CI) HEART score of our study population was 4.6 (4.4 to 4.9). Eighty-five (33%) patients had a low HEART score (0–3), 113 (44%) patients an intermediate HEART score (4–6) and 57 (22%) patients a high HEART score (7–9). Clinical gestalt-based risk of ACS was ‘low’ in 79 (31%) of the patients, ‘intermediate’ in 98 (38%) of the patients and ‘high’ in 78 (31%) patients.

Prevalence of ACS

Seventy-five (29%) patients had an ACS, of which 67 fulfilled the criteria for AMI. In total, 35 patients underwent PCI, 17 patients had a CABG, and 3 patients died within 6 weeks of presentation: one patient due to a massive pulmonary embolus, one due to pneumonia resulting in multiorgan failure with respiratory insufficiency and in one patient the reason of death could not be confirmed, but was suspected to be cardiac dysrhythmia. The risk of ACS in these three patients according to clinical gestalt was low (κ 0.64) for ACS than a low risk based on clinical gestalt. A high HEART score 7 and 30 of 78 patients with a low risk based on gestalt also had a high HEART score, 43 of 98 patients with a moderate risk based on gestalt had a HEART score 4–7 and 30 of 78 patients with a high risk based on gestalt had a HEART score 7–10. Overall diagnostic agreement between HEART score and clinical gestalt was low (κ 0.18).

Diagnostic agreement of clinical gestalt and HEART score

Forty-five of 79 patients with a low risk based on gestalt also had a low HEART score, 43 of 98 patients with a moderate risk based on gestalt had a HEART score 4–7 and 30 of 78 patients with a high risk based on gestalt had a HEART score 7–10. Overall diagnostic agreement between HEART score and clinical gestalt was low (κ 0.18).

Diagnostic value of clinical gestalt and HEART score

Table 4 shows the diagnostic characteristics of HEART score and clinical gestalt including sensitivity, specificity, negative or positive predictive values and likelihood ratios (LR−, LR+). A HEART score ≤3 had a higher sensitivity (93 (85–98)% vs 87 (77–93)% and a lower LR− (0.15 (0.06–0.36) vs 0.35 (0.19–0.64)) for ACS than a low risk based on clinical gestalt. A high HEART score ≥7 on the other hand, had a higher specificity (90 (85–94)% vs 81 (75–87)%) and a higher LR+ (5.2 (3.2–8.5) vs 3.1 (2.2–4.4)) for ACS. Similar results were found for residents and staff, when subgroup analysis was performed, although resident clinical gestalt was slightly better than staff clinical gestalt (table 4).

To make a comparison of overall diagnostic performance, receiver operating curves were plotted for clinical gestalt and for HEART score. In order to be able to compare c-statistics of both...
risk prediction strategies, HEART score was plotted as a 3-category parameter and not as a 10-category parameter (figure 3). Although the c-statistic of HEART score was higher than clinical gestalt (0.81 (0.76–0.86) vs 0.79 (0.73–0.84)), the difference did not reach statistical significance (p=0.13).

DISCUSSION

In our study, we demonstrate that HEART score and clinical gestalt have similar diagnostic accuracy for ACS in a population of patients with undifferentiated chest pain in the ED. This is of clinical relevance, since only a small proportion of patients presenting to the ED with chest pain has or will have an ACS and, therefore, an accurate risk stratification strategy is mandated.11–15 Diagnostic accuracy of HEART score for ACS in our study was comparable with previous studies (c-statistic 0.81 vs 0.83 in a recently published prospective evaluation study).5 However, negative predictive value of a low HEART score (0–3) was lower in our study compared with that latter study (94% vs 98%), whereas positive predictive value of a high HEART score (7–10) was higher (68% vs 50%).4 The higher prevalence of ACS in our population (29% vs 17%) and the higher proportion of high-risk patients (HEART score 7–10; 22% vs 17.5%) might have contributed to these differences. The higher prevalence of ACS in our cohort compared with other populations with patients with undifferentiated chest pain15 16 may be partially attributable to the very accessible primary care system that is in effect in the Netherlands. In our hospital, the vast majority (83%) of cardiac ED patients are referred by their general practitioner, which increases the a priori chance of ACS when visiting the ED. The low negative predictive value of a low HEART score as found in our study may be a concern for any physician wanting to use HEART score for accurately classifying patients as low risk and warrants further investigation in other populations with a high ACS prevalence.

We are aware of only one other study that has compared the diagnostic accuracy of HEART score and gestalt before: Mahler et al17 showed that HEART score had a net reclassification improvement of 10% versus unstructured clinical judgement, mainly due to the identification of more true-negative results. However, HEART score and clinical gestalt comparisons in their study were made in a preselected low-risk population with a TIMI score <2 and not (as in our study) in a more realistic unsellected ED population.

In our study, we found that a low HEART score had a lower LR− for ACS than a low risk based on clinical gestalt and a high HEART score had a higher LR+ than a high risk based on clinical gestalt. Nevertheless, overall diagnostic accuracy of HEART score was equal to the accuracy of clinical gestalt. HEARTS (3),18 a further refinement of HEART score (which uses LR analysis to give appropriate weight to the individual elements of the HEART score), has been shown to have an even better diagnostic accuracy than the original HEART score and it should be anticipated that the difference with clinical gestalt in terms of overall diagnostic accuracy would have been larger when HEARTS(3) instead of HEART would have been used. However, standard ECG and troponin measurements at 2-h

| Table 3 | Prevalence of acute coronary syndrome (ACS) in various risk-groups |
|---------|-----------------------------|-----------------------------|
|          | HEART score | Clinical gestalt |
|          | 0–3 (n=85) | 4–6 (n=113) | 7–10 (n=57) | Low (n=79) | Intermediate (n=98) | High (n=78) |
| ACS (n=75) |               |               |               |               |               |               |
| 5 (6%) | 31 (27%) | 39 (68%) | 10 (13%) | 21 (21%) | 44 (56%) |
| No ACS (n=180) |               |               |               |               |               |               |
| 80 (94%) | 82 (73%) | 18 (32%) | 69 (87%) | 77 (79%) | 34 (44%) |
intervals as in HEARTS(3) are not routinely performed for all patients with chest pain in our hospital and were therefore not a part of our study protocol. Despite that many HEART score elements are also used to determine clinical gestalt-based risk on ACS, overall diagnostic agreement between HEART score and clinical gestalt was low in our study (κ 0.18). This could be explained by the fact that the individual elements that are assigned 0, 1 or 2 points in the HEART score are weighted differently when they are used to determine clinical gestalt-based risk of ACS.

A recently published study by Body et al19 demonstrated that clinical gestalt (based on history and physical exam and rated on a 5-point Likert scale) was insufficient to rule out AMI in the ED, but that sensitivity increased significantly when ECG and troponin levels were also taken into account, as we did in our current study. Surprisingly, resident clinical gestalt performed slightly better than emergency medicine staff clinical gestalt in our study. However, the number of patients seen by staff was small and CIs were consequently wide. Although some literature supports that the more experienced a physician is, the better his recognition skills are and the more accurate clinical gestalt will be,20 this relation should not be overemphasised. Clinical gestalt-based decision-making is riddled with many appreciable errors21 and various studies have recently been published that could not demonstrate a clear relation between experience and clinical gestalt accuracy.22 23

Our study has several shortcomings. First, due to the observational character, the CRF was completed in only a proportion of patients that visited the ED with chest pain and follow-up was incomplete in 10% of the patients. However, comparison of our study population with the total population of patients with chest pain visiting the ED during our study period revealed that patient characteristics and ACS prevalence in our study cohort were an adequate representation of the total population. Furthermore, since this was a monocentre study, the number of participants was limited. Although inherent to our study design, wherein we aimed to compare risk stratification with and without a formal risk stratification tool (HEART), there is a potential for influencing bias, since the same physician scored clinical gestalt and HEART score. More importantly, due to incorporation bias (elements of the respective diagnostic tests under study (HEART score and clinical gestalt) are used to determine the reference standard (ACS)), there might be an overestimation of diagnostic accuracy of both tests. As mentioned earlier, the prevalence of ACS in our cohort was considerably higher than in other undifferentiated chest-pain populations,5 15 16 which makes it difficult to extrapolate our findings to populations with a lower ACS prevalence. Clinical gestalt was represented as a categorical parameter in our study (low, intermediate or high probability of ACS). No consensus exists on how to quantify clinical gestalt and various scoring systems have been used. Since clinical gestalt comprises an immeasurable number of variables, any representation as a categorical parameter might have been an underestimation of the true value of clinical gestalt. Furthermore, follow-up data were obtained from the hospital records. Although there are no hospitals located within a 30 km radius of our hospital and the vast majority of our patients is referred by their primary care physician (regularly to the hospital that they have visited before), there is a chance that ACS has been missed during follow-up. Finally, in our study, we did not look at the respective effects of using HEART score and clinical gestalt on the diagnostic process and treatment disposition of patients with chest pain presenting in the ED. In a recent study, Mahler et al24 demonstrated that

**Table 4** Diagnostic characteristics of HEART score and clinical gestalt

<table>
<thead>
<tr>
<th>Method of low risk estimation</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>NPV (95% CI)</th>
<th>LR− (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEART score ≤3</td>
<td>93 (85 to 98)%</td>
<td>44 (37 to 52)%</td>
<td>94 (87 to 98)%</td>
<td>0.15 (0.06 to 0.36)</td>
</tr>
<tr>
<td>Low risk based on gestalt</td>
<td>87 (77 to 93)%</td>
<td>38 (31 to 46)%</td>
<td>88 (78 to 94)%</td>
<td>0.35 (0.19 to 0.64)</td>
</tr>
<tr>
<td>Subgroup analysis of level of experience:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff (n=77)</td>
<td>81 (62 to 94)%</td>
<td>34 (21 to 49)%</td>
<td>77 (55 to 92)%</td>
<td>0.54 (0.23 to 1.31)</td>
</tr>
<tr>
<td>Residents (n=178)</td>
<td>90 (77 to 97)%</td>
<td>40 (32 to 49)%</td>
<td>91 (81 to 97)%</td>
<td>0.26 (0.11 to 0.61)</td>
</tr>
<tr>
<td>Method of high risk estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEART score 7–10</td>
<td>52 (40 to 64)%</td>
<td>90 (85 to 94)%</td>
<td>68 (54 to 80)%</td>
<td>5.2 (3.2 to 8.5)</td>
</tr>
<tr>
<td>High risk based on gestalt</td>
<td>59 (47 to 70)%</td>
<td>81 (75 to 87)%</td>
<td>56 (44 to 67)%</td>
<td>3.1 (2.2 to 4.4)</td>
</tr>
<tr>
<td>Subgroup analysis of level of experience:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff (n=77)</td>
<td>59 (39 to 78)%</td>
<td>74 (50 to 85)%</td>
<td>55 (35 to 74)%</td>
<td>2.3 (1.3 to 4.0)</td>
</tr>
<tr>
<td>Residents (n=178)</td>
<td>58 (43 to 72)%</td>
<td>84 (76 to 90)%</td>
<td>57 (42 to 71)%</td>
<td>3.6 (2.3 to 5.7)</td>
</tr>
</tbody>
</table>

LR−, negative likelihood ratio; LR+, positive likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

---

**Figure 3** Receiver operating characteristics curves of gestalt an HEART score. Areas under the curve: gestalt: 0.79 (0.73–0.84); HEART score 0.81 (0.76–0.86).
HEART score could substantially reduce cardiac stress testing and cardiac imaging in a population with a low pretest probability of ACS (TIMI<2). Future studies need to be performed to investigate whether this is also true in populations with a higher pretest probability of ACS, like undifferentiated ED populations.

In conclusion, our study demonstrates that 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.

Acknowledgements The authors thank Dr N Veeger of the department of epidemiology and statistics of the University Medical Centre Groningen and the MCL academy for statistic support.

Contributors All performed the data collection and prepared the manuscript. EtA designed the study and coauthored the manuscript. RB and AW assisted in designing the study and reviewed the manuscript.

Competing interests None.

Ethics approval Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
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
Anniek Visser, Albert Wolthuis, Rob Breedveld and Ewoud ter Avest

Emerg Med J 2015 32: 595-600 originally published online September 12, 2014
doi: 10.1136/emermed-2014-203798

Updated information and services can be found at:
http://emj.bmj.com/content/32/8/595

These include:

Supplementary Material
Supplementary material can be found at:
http://emj.bmj.com/content/suppl/2014/09/12/emermed-2014-203798.DC1.html

References
This article cites 24 articles, 7 of which you can access for free at:
http://emj.bmj.com/content/32/8/595#BIBL

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

Topic Collections
Articles on similar topics can be found in the following collections
- Pain (neurology) (969)
- Acute coronary syndromes (192)

Notes

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

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

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