The interobserver agreement of the HEART-score, a multicentre prospective study

Kirsten F. van Meerten[a], Rowan M.A. Haan[a], Ineke M.C. Dekker[b], Henriëtte J.J. van Zweden[a], Erik W. van Zwet[c] and Barbra E. Backus[d]

Background and importance **Chest pain is one of the most common presentations to the emergency department (ED). The HEART-score is used to assess the 30-day risk of developing a major adverse cardiac event (MACE). The HEART-score enables clinicians to classify patients in low, intermediate, or high-risk groups though little is known as to whether this can be done reliably and reproducibly in a prehospital setting.**

Objective **The aim of this study was to compare the interobserver agreement of the HEART-score between ambulance nurses and ED physicians.**

Design, settings, and participants **Patients ≥18 years, with chest pain of suspected cardiac origin presented by ambulance to the EDs of four regional hospitals, were prospectively enrolled between October 2018 and April 2019.**

Outcomes measure and analysis **The primary endpoint was interobserver agreement of the HEART-scores calculated by ambulance nurses compared to those calculated by ED physicians. Agreement was measured using Cohen's Kappa (*K*) both for overall HEART-score and dichotomized HEART categories. A secondary endpoint was the occurrence of a MACE at 30 days after inclusion.**

Main results **A total of 307 patients were enrolled of which 166 patients were male (54%). The mean age was 64.8 years. In 23% (95% confidence interval, 18–27), patients were scored in the low-risk category by both ambulance nurses and ED physicians. The *K* for the overall HEART-score compared between ambulance nurses and ED physicians was 0.514. The *K* for the low-risk category versus intermediate and high-risk category was 0.591. Both are defined as 'moderate'. MACE within 30 days occurred in 64 patients (21%). In the low-risk group as defined by the ambulance nurses, there was a 7% risk of MACE compared to an average 5% MACE risk in the ED physician group.**

Conclusions **The moderate interobserver agreement of the HEART-score does not currently support the use of the HEART-score by ambulance nurses in a prehospital setting. Training for prehospital nurses is vital to ensure that they are able to calculate the HEART-score accurately.**
***European Journal of Emergency Medicine* 28: 111–118**

European Journal of Emergency Medicine 2021, **28**:111–118

Keywords: adult, ambulances, chest pain, cardiology, electrocardiography, humans, observer variation, prospective studies, risk factors, troponin I

[a]Emergency Department, Albert Schweitzer Ziekenhuis Dordrecht & Zwijndrecht, [b]Emergency Department, Franciscus Gasthuis & Vlietland Rotterdam, [c]Department of Medical Statistics, Leids Universitair Medisch Centrum Leiden and [d]Emergency Department, Leids Universitair Medisch Centrum Leiden

Correspondence to Kirsten F. van Meerten, MD, MSc, Emergency Department, Albert Schweitzer Ziekenhuis, Postbus 444, 3300 AK Dordrecht, The Netherlands
Tel: +31613924023; e-mail: meerten@heartscore.nl

Received 14 May 2020 **Accepted** 26 August 2020

Introduction

Chest pain is one of the most common presentations to the emergency department (ED). Whilst it is important to diagnose acute coronary syndrome (ACS) as early as possible so that the correct treatment can be initiated. In approximately 80% of chest pain patients in the ED, no ACS is present [1–3]. Despite the fact that the vast majority of patients will not be experiencing an ACS, many patients are still admitted for observation and may receive unnecessary treatment for an ACS [4]. Research is available which promotes a rapid diagnosis of ACS by developing risk stratification tools for the ED [5,6]. The HEART-score, standing for History, ECG, Age, Risk factors, and Troponin (Fig. 1), is a chest pain risk stratification tool that can help clinicians differentiate patients that are high from those that are low risk for ACS by assessing the 30-day risk of developing a major adverse cardiac event (MACE). It divides patients presenting with chest pain into three risk categories; low-risk (1–3 points), intermediate-risk (4–6 points), and high-risk (7–10 points) [7]. The HEART-score is a well tolerated, easy, and reliable tool that has been validated in multiple studies [3,8–11]. One of the main benefits of the HEART-score is that low-risk patients can be safely discharged from the ED without requiring further investigation as the negative predictive value >98% of MACE risk in the low HEART category was shown to be ≤2% [8–10].

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.euro-emergencymed.com).

0969-9546 Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved. DOI: 10.1097/MEJ.0000000000000758

Fig. 1

History	Highly suspicious	2	
	Moderately suspicious	1	
	Slightly suspicious	0	
ECG	Significant ST-deviation	2	
	Non-specific repolarisation disturbance, LBBB or PMR	1	
	Normal	0	
Age	≥ 65 year	2	
	45 – 65 year	1	
	≤ 45 year	0	
Risk factors	≥ 3 risk factors or history of atherosclerotic disease	2	
	1 or 2 risk factors	1	
	No risk factors known	0	
Troponin	≥ 3x normal limit	2	
	1-3x normal limit	1	
	≤ normal limit	0	
		Total	

HEART-score. LBBB, left bundle branch block ; PMR, pacemaker rhythm.

However, a question that naturally arises is whether low-risk chest pain patients actually need to be seen in the ED at all. The use of a chest pain tool that could decrease ambulance conveyance to the hospital could have the additional benefit of reducing overcrowding and healthcare costs. The current prehospital model in the Netherlands ensures that all ambulances are staffed by a highly trained nurse who could potentially calculate the HEART-score prior to transfer to the hospital and make a decision on whether to convey or not [12].

Two preconditions need to be met in order to demonstrate that ambulance nurses could make decisions on the disposition of chest pain patients. First, they must be able to calculate the HEART-score as accurately as an experienced clinician. Second, it must be shown that the decisions made based upon the calculated HEART-score were, in fact, correct and well tolerated for patients. In 2018, a prospective, cross-sectional study was conducted by Niven *et al*. [13] in the UK demonstrated strong overall interoperator reliability between the HEART-scores performed by various grades of doctors and nurses at the ED. Other studies by Mahler *et al*. [14,15] suggested an acceptable interobserver agreement between doctors in calculating the HEART-score. This is in contrast to a retrospective study by Wu *et al*. [16] that concluded that there is a substantial discordance in HEART-scores between ED physicians and cardiologists. To date; however, no study has been conducted looking at the interoperator reliability of the HEART-score between the ED and a prehospital setting.

The aim of this study was thus to compare the interobserver agreement of the HEART-score between ambulance nurses and ED physicians.

Methods

Study design

This was a prospective, multicentre, study conducted at four regional hospitals and two ambulance services between 1 October 2018 and 7 April 2019. The four regional hospitals were Albert Schweitzer Ziekenhuis, Dordrecht & Zwijndrecht (hospitals 1 and 2), and Franciscus Gasthuis & Vlietland (hospitals 3 and 4). The two ambulance services participating in the study were 'Zuid-Holland Zuid' and 'Rotterdam-Rijnmond'. The study was approved by the local Medical Ethics Review Committee.

In this study, the History, ECG, Age, and Risk factors were calculated by ambulance nurses and ED nurses, ED junior doctors, ED residents, ED physicians or cardiologists prior to receiving the troponin result. These calculations minus the troponin result are referred to

as HEAR(T)-scores in this article. Level of experience between the nurses and doctors was not taken into account.

Inclusion and exclusion criteria

All adults aged ≥18 years with nontraumatic chest pain of suspected cardiac origin who presented by ambulance to the ED and had at least three, separate HEAR(T)-scores calculated of which one had to be calculated by an ambulance nurse, were included. Patients with chest pain that presented to the ED without ambulance and patients who were transported directly to the catheter laboratory because of ST elevation myocardial infarction, were excluded. Patients for whom forms were incomplete to calculate their HEAR(T)-score and patients who refused to participate were also excluded.

Study protocol/data collection

Patients were recruited by consecutive eligibility. The study sets consisted of five identical Case Record Forms (CRF, Supplemental Digital Content 1, supplement digital content 1, *http://links.lww.com/EJEM/A288*) and one Patient Information Form (PIF, Supplemental Digital Content 2, supplement digital content 1, *http://links.lww.com/EJEM/A288*). These sets were available at the ambulance entrance to the ED. Upon arrival at the ED with an eligible patient, ambulance-staff were requested to calculate the HEAR(T)-score. During subsequent ED admission, the HEAR(T)-score was separately calculated by an ED nurse, ED junior doctor, ED resident, ED physician, or cardiologist who were blinded to the other results. Calculations were completed prior to the final diagnosis. Due to the delay in obtaining the troponin result, the scoring for this element was added by the researchers at a later stage and was based on the first troponin performed in the ED, irrespective of time of onset of chest pain.

All four hospital used a high sensitive troponin I assay (hs-cTnI): hospitals 1 and 2 used the Siemens Dimension Vista and the hospitals 3 and 4 employed the Abbott Architect Stat assay. The 99th percentile cutoff values were locally validated and sex-specific based upon the recommendations of the manufacturer. Troponin values between the 99th percentile and three times the clinical cutoff value were scored as '1'; troponin values above three times the clinical cutoff value were scored as '2' and those below the 99th percentile cutoff were scored as '0'.

Training of staff and patient information

A face-to-face teaching presentation was provided to the observers prior to the onset of the study. The presentation contained information about the principles of the HEART-score, a detailed explanation of how to calculate it as well as information about the study itself. Those who could not attend a teaching session were informed by means of e-mail. A specific video was created for this group accessible on https://www.youtube.com/watch?v=x4ubZhqF9tw&t=18s [17], which reinforced the learning from the face-to-face sessions. Finally, all nurses and doctors involved in the study were provided with a study folder (folder, Supplemental Digital Content 3, supplement digital content 1, *http://links.lww.com/EJEM/A288*) and a HEART-score pocket card.

Endpoints

The primary endpoint was the interobserver agreement of the HEART-score as calculated between various grades of nurse and doctor.

The secondary endpoint was the occurrence of MACE within 30 days after inclusion. MACE was defined as occurrence of acute myocardial infarction (AMI), percutaneous coronary intervention (PCI), coronary artery bypass grafting (CABG) or death.

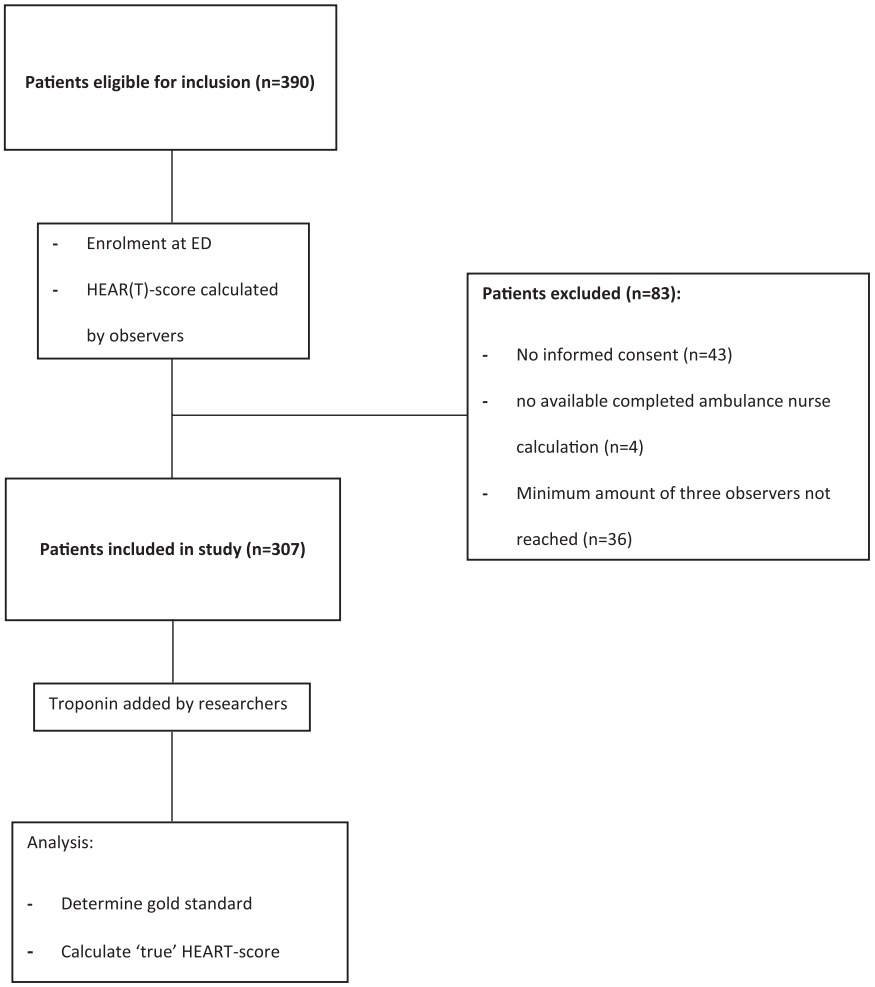
Statistical methods

Descriptive statistics were used to describe the demographic and clinical attributes of the population. To determine the reliability between the ambulance nurses and the ED physicians, the 'gold standard' was used. The gold standard, or clinically relevant comparator, was defined retrospectively as the HEAR(T)-score calculated by the most senior doctor who assessed the patient variously, and in descending order: cardiologist, ED physician, ED resident in training, and ED junior doctor not in training. This is in concordance with clinical practice. The HEART-scores calculated were then further subcategorized into three groups: low, intermediate, and high risk. The agreement between these subgroups was then, in turn, used to calculate the reliability statistics. We examined if patients were allocated to a different risk category by ambulance nurses versus the gold standard, to assess the safety and feasibility of a prehospital HEART-score. As the low-risk group was of greatest practical interest to this study, the HEART-score was dichotomized and statistical analysis was focused on low compared to intermediate and high-risk categories calculated by ambulance nurses versus the gold standard. The reliability between the ED nurses and the gold standard is also determined.

Interobserver agreement is defined as the measure of agreement or reliability between two or more observers when observing the same subject. In this study, one subject was observed by three or more clinicians (nurses and doctors) of differing grades and levels of experience. With a kappa, the degree of agreement was described in one number [18]. To determine the agreement between the ambulance nurses and ED nurses respectively and the gold standard, the Cohen's Kappa (K) was used (Table I, Supplemental Digital Content 4, supplement digital content 1, *http://links.lww.com/EJEM/A288*). Confidence interval (CI) was calculated afterwards.

A convenience sample of 300 patients was included. The data are analyzed using IBM SPSS Statistics, version 24.

Fig. 2



Flow diagram. ED, emergency department.

Results

A total of 390 patients were considered for inclusion (Fig. 2). Eighty-three patients were excluded because of lack of informed consent ($n=43$), no available completed ambulance nurse calculation ($n=4$) and minimum amount of three observers not reached ($n=36$). Thus 307 complete cases with at least three observers per patient were obtained. Of the 307 patients, 166 patients were male (54%) and 141 female (46%). The mean age was 64.8 years (range 19–91, SD ±14.1) (Table 1). During busy hours inclusions appear to have been missed because the ambulance form was not readily available and ambulance nurses already left the ED.

The distribution of the observers was as follows: 307 patients were scored by ambulance nurses, of which 302 patients were also scored by ED nurses, 240 patients by ED residents (residents in training $n=77$, junior doctors not in training $n=163$), 106 patients by ED physicians, and 23 patients were scored by cardiologists. The median HEART-score of the total study population was 5.0 and the mean score was 4.55 (SD ±2.046).

Table 1 Patient characteristics at baseline

Characteristics	Total study population ($n=307$)
Demographic	
Male, *n* (%)	166 (54)
Mean age, years (SD)	64.8 (14)
Cardiac risk factors, *n* (%)	
Hypercholesterolemia	96 (31)
Hypertension	161 (52)
Diabetes mellitus	63 (21)
Current smoking	87 (28)
Positive family history	130 (42)
Obesity (BMI>30 kg/m^2)	68 (22)
History of cardiovascular disease, *n* (%)	126 (41)
Hospital admission, *n* (%)	135 (44)

Table 2 HEART-score category – ambulance nurse versus gold standard

	Total HEART-score ambulance nurse			
Total HEART-score Gold	Low risk	Intermediate risk	High risk	Total
Low risk	69	31	1	101
Intermediate risk	22	119	23	164
High risk	0	12	30	42
Total	91	162	54	307

Table 3 True HEART-score versus major adverse cardiac event

Score	Patients (%)	MACE/*n*	MACE (%)	Death (%)
0–3	29	1/90	1	0
4–6	55	33/170	19	1.3
7–10	15	30/47	64	1.0

MACE, major adverse cardiac event.

Primary endpoint – interobserver agreement

Ambulance nurses versus gold standard

The distribution of the gold standard is shown in Table II (Supplemental Digital Content 5, supplement digital content 1, *http://links.lww.com/EJEM/A288*, gold standard distribution). In 30% (95% CI, 25–35) and 33% (95% CI, 28–38), respectively, the ambulance nurses and gold standard scored patients in a low-risk group (Table 2). In 23% (95% CI, 18–27), patients were scored in the low-risk category by both ambulance nurses and the gold standard. In 11% (95% CI, 7–14) of the cases, the ambulance nurses scored patients in a lower category than the gold standard, in 71% (95% CI, 66–76) the ambulance nurses scored the patients in the same category and in 18% (95% CI, 14–22) in a higher category. In 7% (95% CI, 4–10), ambulance nurses scored patients in a low-risk category whereas the gold standard had placed these patients in an intermediate or high-risk category. A MACE occurred in 3 of these 22 cases.

The Cohen's Kappa (K) for the overall HEART-score between the ambulance nurses and the gold standard was 0.514 (95% CI, 0.426–0.602), defined as 'moderate' (Table III, Supplemental Digital Content 6, supplement digital content 1, *http://links.lww.com/EJEM/A288*). When comparing the low-risk to intermediate and high-risk categories in these two groups, the K is 0.591 (95% CI, 0.491–0.691); also defined as 'moderate'.

The agreement between ED nurses and the gold standard is shown in Supplemental Digital Content 7 supplement digital content 1, *http://links.lww.com/EJEM/A288* (text, Tables IV and V).

Secondary endpoint

Major adverse cardiac event

To assess the occurrence of MACE versus the HEART-score, a so-called 'true' HEART-score was calculated by the researchers with >2 years of experience in calculating a HEART-score. This calculation was done retrospectively based on the available ED documents and ECG, blinded for the outcome. The results of MACE within 30 days are given in Table 3 (see also Figure I, Supplemental Digital Content 8, graph, supplement digital content 1, *http://links.lww.com/EJEM/A288).*

A total of 64 patients (21%) reached a MACE within 30 days. Of the patients who reached a MACE 51 patients had an AMI (80%), 53 underwent coronary angiography (83%), 38 underwent a PCI (59%), 7 underwent CABG (11%), and 7 patients died (11%) within 30 days. When using the 'true' HEART-score, there was a 1% MACE risk in the low-risk HEART category.

However, in the patients with a low HEART-score as defined by the ambulance nurses there was a 7% (6/91) risk of MACE. A similar figure was seen in the low HEART-scores of the ED nurses with a MACE rate of 5% (5/100) in the low-risk category (Table 4) and for junior doctors not in training with a MACE rate of 10% (5/52). For senior residents and cardiologists, the MACE rate was 0% and for ED physicians 3% (1/36) for those patients scored in the low-risk category. This is an average MACE risk of 5% in the low-risk category scored by the ED physicians.

All cases were retrospectively reviewed in detail where it was found that the principle areas of poor agreement included the 'History' and 'Risk factors' elements. This resulted in a trend towards an underestimation of the HEAR(T)-scores calculated by ambulance nurses, ED nurses and junior doctors in patients who subsequently suffered a MACE.

Discussion

The aim of this study was to investigate if the HEART-scores calculated by ambulance nurses were comparable with HEART-scores calculated by ED physicians. The principal finding of a Cohen's Kappa (*K*) of 0.514 defined as 'moderate' for the overall HEART-score suggests that ambulance nurses do not calculate the HEART-score as accurately as the gold standard. This is also seen when dichotomizing the HEART-score and comparing the low-risk category to the intermediate and high-risk group ($K = 0.591$). When comparing the ED nurses to the gold standard, the *K* is 0.603, also defined as 'moderate'.

In this study, we were mainly interested in cases where ambulance nurses scored patients in a different risk

Table 4 Major adverse cardiac event rate per HEART category

HEART category versus MACE	HEART category	*N* (total)	MACE	MACE/*n*	%
Ambulance nurse	0–3	91	6	6/91	7
n=307	4–6	162	29	29/162	18
	7–10	54	29	29/54	54
Total		307	64		
ED nurse	0–3	100	5	5/100	5
N=302	4–6	151	30	30/151	20
	7–10	51	28	28/51	55
Total		302	63		
Junior doctor not in training	0–3	52	5	5/52	10
N=163	4–6	90	16	16/90	18
	7–10	21	11	11/21	52
Total		163	32		
Resident in training	0–3	24	0	0/24	0
N=77	4–6	43	9	9/43	21
	7–10	10	7	7/10	70
Total		77	16		
ED physician	0–3	36	1	1/36	3
N=106	4–6	54	13	13/54	24
	7–10	16	12	12/16	75
Total		106	26		
Cardiologist	0–3	7	0	0/7	0
N=23	4–6	12	2	2/12	17
	7–10	4	3	3/4	75
Total		23	5		

ED, emergency department; MACE, major adverse cardiac event.

category from ED doctors. This was in order to assess the feasibility of implementing a system whereby the HEART-score could be used in a prehospital environment to help decide whether patients should be conveyed to hospital or managed in the community. Overall, ambulance nurses placed the patients in a higher risk category than their ED doctor counterparts (18%). This trend of 'over-scoring' patients is preferable than underestimating the score from a safety perspective. It is favorable to have more false positives resulting in a potential unnecessary visit to the ED, than false negatives with the risk of missing patients with myocardial ischemia.

In 22 cases (7%), the ambulance nurses scored the patients in a low-risk category where the gold standard scored these patients in an intermediate category. In these cases, the ambulance nurses could have chosen to leave the patient at home based on the HEART-score. In 3 of these 22 cases, the patient developed a MACE within 30 days, none of these patients died. Analysis of the data revealed that this misclassification was likely due to an underestimation of the elements History and Risk factors which had Cohen's Kappa (*K*) values of 0.171 (History) and 0.417 (Risk factors), defined as 'poor' and 'moderate', respectively. The trend of poor agreement for History in contrast to other elements of the HEART-score is seen in earlier studies [13,19]. The accuracy of chest pain history is influenced by many factors including, but not limited to, the level of training and experience of the (para)medic taking the history. The unreliability of the history or clinical gestalt has been illustrated in studies by Oliver and Carlton *et al*. [20,21]. Lack of training and an understanding of what constitutes a 'Risk factor' may also be responsible for the poor agreement observed for this element of the HEART-score.

In regards to the secondary endpoint, the percentage of patients reaching MACE within 30 days after inclusion in this study is higher than in earlier studies concerning the HEART-score conducted by Backus *et al*. [10,11]. An explanation for this observation is that the population of chest pain patients presenting to the ED by ambulance is a higher risk group than walk-in patients presenting with chest pain to the ED, for example, self-referrals or patients referred by a general practitioner.

In both hospitals, hs-cTn is used where at least one value above the 99th percentile of the upper reference limit is scored as 1 or 2. Although the introduction of hs-cTn has increased the early identification of (small) NSTEMIs, compared to the studies performed almost 10 years ago the percentages of PCI, CABG, and death are similar or even lower [10,11]. Although there are several articles that have explored the use of hs-cTn in conjunction with the HEART-score with differential scoring for the result based upon not only the 99th percentile but also the so-called 'limit of detection' strategy [8] this has not been formally and prospectively validated. Research is currently underway that will hopefully clarify this [22].

Implementation in clinical practice

Our findings suggest that the HEART-score cannot, at this stage, be performed by ambulance nurses with similar accuracy as those calculated by ED physicians. Whilst the principle remains that an accurately calculated HEART-score in the prehospital setting could potentially result in fewer conveyances to hospital; a higher level of training and potentially certification in the calculation of the HEART-score by prehospital nurses will be required before this aspiration could be realized. Prehospital nurses receive a high level of training in the Netherlands and it

seems reasonable that given this fact, a training and certification system could be designed that would improve their ability to accurately perform the HEART-score.

An additional, logistical problem is the availability of troponin testing in the prehospital environment. This could potentially be addressed by access to a Point-Of-Care Troponin Test (POC-Troponin). Certain POC-Troponin assays now commercially available are as accurate as laboratory-based assays [23,24]. There are current feasibility studies underway investigating the use of POC-Troponin in ambulances [25,26].

Study limitations

In this study, the level of clinical experience and training of the multiple groups of observers was not taken into account. However, Niven *et al.* [13] confirm that ED nurses and ED doctors can reliably calculate the HEART-score irrespective of grade and experience. The effect of the level of grade of observers in this study is, therefore, thought not to be significant. But proper education in the HEART-score itself and its use in clinical was needed before the start of the study. It is; however, difficult to quantify the effectiveness of the training methods we employed given the absence of a control group.

Originally, we planned to study the interobserver agreement of the HEAR(T)-score between ambulance nurses on the one hand, and ED junior doctors, ED residents, ED physicians and cardiologists on the other. Of course, not every patient is evaluated by all 'types' of physicians. To ensure a sufficient number of examples of each type, we decided to collect data on 300 patients. We have performed the analyses as intended, but our primary focus shifted to comparing the HEART-score of the ambulance nurses to the HEART-score as determined by the most senior attending doctor, defined as gold standard. We believe that this is the most relevant comparator because it best reflects the HEART-score as it is determined in clinical practice. Now the planned sample size of $n = 300$ is large enough to estimate the proportion of patients where the ambulance nurse agrees with the most senior attending doctor with sufficient accuracy. For example, $n = 300$ ensures that the width of the 95% CI is ±0.05 when the estimated proportion is 0.25 or 0.75.

There is a possibility of selection bias in this study. Despite the aspiration of consecutive eligibility for this study, during busy periods it was harder to ensure that ambulance nurses were completing the forms and HEART-score calculations necessary for the study. The cohort of patients that went directly to the catheter laboratory was also not included in this study.

A strength of our study is that it is unique. We compared prehospital calculations of the HEART-score with the HEART-scores calculated by ED physicians. No earlier studies regarding this comparison are known to us.

Conclusion

In this study, the interobserver agreement of the HEART-score calculated by ambulance nurses compared to the HEART-score calculated by ED physicians is moderate ($K = 0.514$). Our study does not currently support the use of the HEART-score by ambulance nurses in a prehospital setting. The implementation of standardized and comprehensive training system on the HEART-score for prehospital clinicians coupled with reliable POC troponin testing is required if this conclusion is to be challenged in the future.

Acknowledgements

We would like to thank Timo Bevelander (Ambulance Department, Zuid-Holland Zuid) and Nancy van der Waarden (Ambulance Department, Rotterdam-Rijnmond) for their help and enthusiasm during the set-up and implementation of this study. We would like to thank William Niven for his time, for the positive feedback and helpful comments for correction or modification.

Conflicts of interest

There are no conflicts of interest.

References

1 Lee TH, Goldman L. Evaluation of the patient with acute chest pain. *N Engl J Med* 2000; **342**:1187–1195.
2 Hollander JE. Risk stratification of emergency department patients with chest pain: the need for standardized reporting guidelines. *Ann Emerg Med* 2004; **43**:68–70.
3 Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Heart J* 2008; **16**:191–196.
4 Cotterill PG, Deb P, Shrank WH, Pines JM. Variation in chest pain emergency department admission rates and acute myocardial infarction and death within 30 days in the Medicare population. *Acad Emerg Med* 2015; **22**:955–964.
5 Body R. Incentivising the rapid diagnosis of acute coronary syndromes in England and Wales. *Eur J Emerg Med* 2020; **27**:237–238.
6 Hausfater P, Velly L. Diagnosis of acute coronary syndrome in the emergency room: the dictatorship of high-sensitivity cardiac troponin and major adverse cardiac events. *Eur J Emerg Med* 2020; **27**:239–240.
7 The HEART score official website. Available from: www.heartscore.nl.
8 Poldervaart JM, Reitsma JB, Backus BE, Koffijberg H, Veldkamp RF, Ten Haaf ME, *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.
9 Six AJ, Cullen L, Backus BE, Greenslade J, Parsonage W, Aldous S, *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.
10 Backus BE, Six AJ, Kelder JC, Bosschaert MA, Mast EG, Mosterd A, *et al.* A prospective validation of the HEART score for chest pain patients at the emergency department. *Int J Cardiol* 2013; **168**:2153–2158.
11 Backus BE, Six AJ, Kelder JC, Mast TP, van den Akker F, Mast EG, *et al.* Chest pain in the emergency room: a multicenter validation of the HEART score. *Crit Pathw Cardiol* 2010; **9**:164–169.
12 Backus BE, Tolsma RT, Boogers MJ. The new era of chest pain evaluation in the Netherlands. *Eur J Emerg Med* 2020; **27**:243–244.
13 Niven WGP, Wilson D, Goodacre S, Robertson A, Green SJ, Harris T. Do all HEART scores beat the same: evaluating the interoperator reliability of the HEART score. *Emerg Med J* 2018; **35**:732–738.
14 Mahler SA, Hiestand BC, Goff DC Jr, Hoekstra JW, Miller CD. 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.
15 Mahler SA, Riley RF, Hiestand BC, Russell GB, Hoekstra JW, Lefebvre CW, *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.
16 Wu WK, Yiadom MY, Collins SP, Self WH, Monahan K. Documentation of HEART score discordance between emergency physician and cardiologist evaluations of ED patients with chest pain. *Am J Emerg Med* 2017; **35**:132–135.
17 Nederlandse introductie in de HEART-score. Available from: https://www.youtube.com/watch?v=x4ubZhqF9tw&t=18s.
18 Nelson KP, Edwards D. Measures of agreement between many raters for ordinal classifications. *Stat Med* 2015; **34**:3116–3132.
19 Oliver JJ, Streitz MJ, Hyams JM, Wood RM, Maksimenko YM, Long B, *et al*. An external validation of the HEART pathway among emergency department patients with chest pain. *Intern Emerg Med* 2018; **13**:1249–1255.
20 Oliver G, Reynard C, Morris N, Body R. Can emergency physician gestalt "rule in" or "rule out" acute coronary syndrome: validation in a multicenter prospective diagnostic cohort study. *Acad Emerg Med* 2020; **27**:24–30.
21 Carlton EW, Than M, Cullen L, Khattab A, Greaves K. 'Chest pain typicality' in suspected acute coronary syndromes and the impact of clinical experience. *Am J Med* 2015; **128**:1109–1116.e2.
22 Khand A, Frost F, Chew P, Fisher M, Mullen L, Grainger R, *et al*. Modified heart score improves early, safe discharge for suspected acute coronary syndromes: a prospective cohort study with recalibration of risk scores to undetectable high sensitivity troponin T limits. *J Am Coll Cardiol* 2017; **69**:238.
23 Juliano M, Wason C. Comparison of point-of-care versus laboratory troponin testing in an emergency department setting. *Mil Med* 2017; **182**:e1938–e1940.
24 Pickering JW, Young JM, George PM, Watson AS, Aldous SJ, Troughton RW, *et al*. Validity of a novel point-of-care troponin assay for single-test rule-out of acute myocardial infarction. *JAMA Cardiol* 2018; **3**:1108–1112.
25 Ishak M, Ali D, Fokkert MJ, Slingerland RJ, Tolsma RT, Badings E, *et al*. Fast assessment and management of chest pain patients without ST-elevation in the pre-hospital gateway (FamouS Triage): ruling out a myocardial infarction at home with the modified HEART score. *Eur Heart J Acute Cardiovasc Care* 2018; **7**:102–110.
26 Kip MMA, Koffijberg H, Moesker MJ, IJzerman MJ, Kusters R. The cost-utility of point-of-care troponin testing to diagnose acute coronary syndrome in primary care. *BMC Cardiovasc Disord* 2017; **17**:213.