

BMJ Open Performance of risk scores for coronary artery disease: a retrospective cohort study of patients with chest pain in urgent primary care

Michelle Kleton,¹ Amy Manten ,¹ Iris Smits,¹ Remco Rietveld,² Wim A M Lucassen,¹ Ralf E Harskamp¹

To cite: Kleton M, Manten A, Smits I, *et al*. Performance of risk scores for coronary artery disease: a retrospective cohort study of patients with chest pain in urgent primary care. *BMJ Open* 2021;**11**:e045387. doi:10.1136/bmjopen-2020-045387

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045387>).

MK and AM are joint first authors.

Received 06 October 2020
Accepted 14 November 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of General Practice, Amsterdam Public Health Research Institute, Amsterdam UMC - University of Amsterdam, Amsterdam, The Netherlands
²Huisartsenpost, Huisartsenorganisatie Noord-Kennemerland (HONK), Alkmaar, The Netherlands

Correspondence to

Dr Ralf E Harskamp;
r.e.harskamp@amsterdamumc.nl

ABSTRACT

Objective To evaluate the diagnostic performance of the Marburg Heart Score (MHS), INTERCHEST, Gencer rule, Bruins Slot rule and compare these with unaided clinical judgement in patients with chest pain in urgent primary care.

Design Retrospective, cohort study.

Setting Regional primary care facility responsible for out-of-hours primary care for a quarter-million people in the Netherlands.

Participants Consecutive patients aged ≥ 18 years who were evaluated for chest pain.

Main outcome measures Discriminatory ability (C-statistic), sensitivity, specificity, positive and negative predictive values (PPV/NPV). The reference standard involved a composite endpoint of the occurrence of death, acute coronary syndrome or coronary revascularisation (=major adverse cardiac events; MACE) up to 6 weeks after initial contact.

Results A total of 664 patients were included, of whom 4.8% (n=32) had a MACE event. C-statistics for MHS, INTERCHEST, Gencer and Bruins Slot rule were: 0.77 (95% CI 0.69 to 0.84), 0.85 (95% CI 0.78 to 0.92), 0.72 (95% CI 0.63 to 0.81) and 0.72 (95% CI 0.63 to 0.81), respectively. Optimal diagnostic accuracy was found for MHS ≥ 2 (sensitivity=81.3%, specificity=67.1%, PPV=11.1%, NPV=98.6%), INTERCHEST ≥ 2 (sensitivity=87.5%, specificity=78.8%, PPV=17.3%, NPV=99.1%), Gencer ≥ 2 (sensitivity=84.4%, specificity=37.8%, PPV=6.4%, NPV=98.0%) and Bruins Slot ≥ 2 (sensitivity=90.6%, specificity=40.8%, PPV=7.2%, NPV=98.9%). Physicians referred 157 patients (23.6%) and missed 6 out of 32 MACEs (sensitivity=81.3%, specificity=79.3%, PPV=16.6%, NPV=98.8%). Using INTERCHEST with a referral threshold of ≥ 2 points, 4 MACEs would have been missed and 162 patients (24.4%) referred. The other risk scores resulted in far higher referral rates.

Conclusion While available risk scores have reasonable to good discriminatory properties, they do not outperform unaided clinical judgment for evaluating chest pain in urgent primary care. Only the INTERCHEST score may slightly improve risk stratification.

INTRODUCTION

Chest pain is a common symptom for consulting a general practitioner (GP),

Strengths and limitations of this study

- First head-to-head evaluation of four clinical decision aids in (out-of-office-hours) primary care.
- Representative patient sample with low percentage of lost to follow-up.
- Evaluation of both cardiovascular and non-cardiovascular major events.
- Retrospective design.
- Limited number of major adverse cardiac events.

accounting for approximately 0.7%–3.0% of consultations.^{1 2} Of all patients presenting with chest pain in primary care, up to 16.0% is diagnosed with cardiovascular disease and 1.5%–3.6% with acute coronary syndrome (ACS).³ One of the critical tasks of GPs is to distinguish possible life-threatening causes of chest pain from far more common non-severe conditions. Making this distinction can at times be challenging, particularly when a patient presents with less clear-cut symptoms.

GPs rely on history taking, physical examination, (prior) experience and gut feeling in their attempt to distinguish between acute and non-acute pathology.⁴ Prior research has shown that GPs would be very receptive for having a reliable clinical risk score at their disposal as a clinical decision aid.⁵ A number of clinical risk scores (Marburg Heart Score (MHS), INTERCHEST score, Gencer rule, Bruins Slot rule) have been developed in primary care over the past decade. Of those, only the Bruins Slot rule was developed for ruling out ACS, whereas the others were developed in patient populations with predominantly stable coronary artery disease. In this study, we aimed to evaluate the diagnostic accuracy of these scores for chest pain evaluation in an urgent primary care setting in which we focused on major adverse cardiac events (MACE). Moreover, we evaluated whether these risk scores may actually improve

the diagnostic performance of unaided clinical judgement of the GP.

METHODS

We reported this diagnostic accuracy study in accordance with the Standards for Reporting of Diagnostic Accuracy Studies 2015 statement.⁶ The study protocol was evaluated by our institution's Medical Ethical Review Committee, and was granted a waiver because of its observational nature.⁷

Study design

This study involved a retrospective, observational cohort of consecutive patients evaluated for chest pain at a large regional out-of-hours primary care facility in Alkmaar, The Netherlands. The facility is responsible for out-of-hours urgent primary care for 245 000 inhabitants. Eligible patients were aged 18 years or older and were evaluated in person for chest pain by a GP between 1 January 2017 and 31 December 2017. Evaluation for clinical outcomes were obtained out to August 2019.

Background setting of the study

In the Dutch healthcare system, patients with symptoms of chest pain during evenings or weekends typically contact a centralised emergency number or a centralised out-of-hours primary care facility. In both instances, telephone triage will follow. The triage assistant differentiates between high versus low(er) probability of a potential serious underlying condition using a standardised digital triage tool. In a substantial number of patients, triage results in direct ambulance activation, which typically involves patients with signs of haemodynamic compromise and/or classical high-risk symptoms and risk factors. In the remaining cases, triage assistants decide on the most fitting action. In some cases telephone advice is sufficient, while in others the GP will evaluate the condition of the patient during a face-to-face consultation using clinical appearance, symptom presentation, physical examination and sometimes additional testing (eg, ECG or laboratory tests). This can occur at the patient's home, or more often at the primary care facility.

Data collection

Data were collected and processed using a secure electronic data capturing platform (Castor EDC, Ciwit BV, Amsterdam, The Netherlands).⁸ Obtained data involved baseline information and subsequent follow-up. Baseline data included; demographics, medical history, relevant use of medication, anamnesis, physical examination, additional diagnostic workup if available (ie, resting 12-lead ECG), working diagnosis and recommended course of action. The working diagnosis illustrates the most probable diagnosis and is registered during GP consultation as an International Classification of Primary Care (ICPC) code. It, therefore, reflects the GPs assessment of the complaint.

Every Dutch citizen is registered with a GP who holds an electronic health record for each patient. These records include the correspondence from other healthcare services, such as admission or discharge notes from the emergency department (ED)/hospital, as well as outpatient letters from hospital-based physicians. Data available from this electronic health record were used to obtain follow-up information on relevant cardiovascular events and other major events linked to the initial complaint of chest pain. Final diagnosis were also registered as ICPC codes and were all adjudicated by a panel of experts. Clinical follow-up data were completed in August 2019. The process of data collection was published previously in greater detail.⁷

Calculation of risk scores

As shown in [table 1](#), the elements used in the Marburg Heart, INTERCHEST, Gencer and/or Bruins Slot rules include features such as; age and sex, a known history of cardiovascular disease, increasing pain with exercise, reproducibility of the pain by palpation, duration and location of the pain. Some of the rules also contain more subjective elements, reflecting whether the patient or physician suspects an underlying cardiac condition. For the feature 'patient assumes a cardiac origin', we looked for cues of the patient's concern in the consultation note. Whether the GP suspected a cardiac condition was determined by: (1) the presence of statements of concern in the medical record, (2) the working diagnosis/differential diagnosis and/or (3) the subsequent management action following assessment. Finally, the Bruins Slot rule includes radiation of pain and presence of nausea/sweating. The definition of pain radiance was not specified in the original publication, and we interpreted this as the presence of typical cardiac radiation of chest pain, similar to what is used in the HEART score.⁹ In consideration of the fact that each GP-consultation varies from the next, we presumed the absence of a score's element when such a feature was not recorded in the electronic health record. Calculation of risk scores occurred after completion of data collection and the responsible investigator was blinded for clinical outcomes. We assessed the diagnostic accuracy for each clinical risk score with multiple cut-off points. A score exceeding the cut-off value would be considered a high-risk case in which immediate referral would be warranted.

Definition of unaided GP decision making

We defined unaided GP decision making as the process of GP assessment and subsequent management (others may refer to this as 'clinical gestalt'). This process can vary, depending on the amount of information the GP acquires; in some cases history taking and clinical evaluation might be sufficient, while others warrant additional testing (ie, resting 12-lead ECG) or discussion with a (hospital-based) colleague. The electronic health record provides detailed information on this process and subsequent actions. Cases which were referred to the

Table 1 Components of the clinical decision rules

Marburg heart score¹¹	
Known clinical vascular disease*	1
Age/sex (F ≥65 years or M ≥55 years)	1
Increased pain with exercise	1
Pain not reproducible by palpation	1
Patient assumes pain is of cardiac origin	1
INTERCHEST¹⁶	
History of CAD†	1
Age/sex (F ≥65 years or M ≥55 years)	1
Increased pain with exercise	1
Pain reproducible by palpation	-1
Physician suspects serious diagnosis	1
Pain feels like 'pressure'	1
Genger rule²¹	
History of CVD*	2
Age/sex (F ≥65 years or M ≥55 years)	2
Increased pain with exercise	1
Pain not reproducible by palpation	1
CVD risk factor‡	2
Duration of pain 1–60 min	1
Substernal location of pain	1
Bruins slot rule²³	
History of CAD†	2
Male sex	5
Presence of radiation of pain§	8
Presence of nausea/sweating	5

*History of myocardial infarction, transient ischaemic attack, cerebrovascular accident, peripheral artery disease, previous percutaneous coronary intervention or coronary artery bypass graft.

†History of myocardial infarction, previous percutaneous coronary intervention or coronary artery bypass graft.

‡Family history of CVD, diabetes mellitus, hypertension, hypercholesterolaemia, smoking or obesity.

§Radiation of pain was deemed present in case of typical cardiac radiation, meaning; radiation to jaw(s), left shoulder or arms as is used in the HEART score.

CAD, coronary artery disease; CVD, cardiovascular disease.

ED/hospital by the GP were considered to be high risk, whereas cases which were not referred were considered low risk for a possible major (cardiovascular) event. Thus, we evaluated unaided GP decision-making dependent on the course of action after evaluation.

Clinical outcomes

Primary outcome: MACE

The primary outcome was the occurrence of MACE within 6 weeks of initial contact. MACE was defined as a composite consisting of death from any cause, ACS or coronary revascularisation. The 6 week observation period was derived from previous research validating

the HEART score, and allows for observing delayed-type events related to the initial presentation.¹⁰

Secondary outcome: coronary artery disease

A secondary outcome was the diagnosis of (any) coronary artery disease (CAD) up to 6 months after initial contact, which we defined as the occurrence of MACE, as well as a final diagnosis of cardiac related pain or pressure (ICPC codes: K01, K02, K03), stable angina pectoris (ICPC: K74), ischaemic heart disease (ICPC: 76), or ischaemic cardiomyopathy, defined as ischaemic heart disease plus heart failure (ICPC: K77). CAD includes both recurrent episodes of existing disease, as well as new (acute) diagnoses. The 6 months follow-up for the occurrence of CAD is in accordance with the derivation studies of the clinical decision rules.

Secondary outcome: major event

This study was designed to include all patients presenting with chest pain. Since chest pain can be an indicator for cardiovascular as well as serious non-cardiovascular diseases, we included an assessment of all major events occurring within 6 weeks after initial contact. This composite of 'major event' is summarised in table 2 and includes diagnoses such as pulmonary embolisms, severe congestive heart failure and cerebrovascular accidents.

Statistical methods

We expressed diagnostic accuracy for each of the risk scores for detecting (1) 6-week MACE, and (2) the secondary outcomes (6-month CAD and 6-week major event) at optimal thresholds using sensitivity, specificity, positive and negative predictive values (PPV/NPV), with 95% CIs. We displayed the overall discriminatory properties using C-statistics.

Patient and public involvement

For this study, we involved triage assistants and affiliated GPs of the out-of-hours primary care facility in Alkmaar, the Netherlands. The experience and priorities of these parties supported the investigators in the design of the study and in the development of research questions. The Dutch Heart Foundation and GPs within our academic network were informed prior to the start of the project. This study did not include direct patient involvement. However, after peer review and acceptance for publication, we will share our findings with relevant patient organisations.

RESULTS

Patient characteristics

GPs evaluated a total of 770 patients for chest pain. From these we excluded 83 patients as an objection was raised to share medical data for research purposes (as a result of the instalment of new European privacy regulations). Of the remaining patients, we could not obtain follow-up information in 23 (3.3%) patients, which left us with a study population of 664 patients. A

Table 2 Definition of major event versus non-major event (cut-off <6 weeks)

	Final diagnosis	Management
Major event	Death from any cause	
	Acute coronary syndrome	
	Urgent coronary revascularisation	
	Pulmonary embolism	
	Thoracic aortic aneurysm (dissection or ruptured)	
	Severe/acute congestive heart failure	Hospitalisation
	Severe peri(myo)carditis	Hospitalisation
	(Tension) pneumothorax	Hospitalisation
	Severe pneumonia	Hospitalisation
	Symptomatic atrial fibrillation	Hospitalisation (cardioversion or converted through medication)
	Aortic valve stenosis	Hospitalisation
	CVA /TIA	
	Traumatic event (with significant impact)	Hospitalisation
	Inflammatory processes such as appendicitis, pancreatitis, cholecystitis	Hospitalisation
	Other, such as: exacerbation COPD or hypertensive crisis	
Non-major event	Mild congestive heart failure	Outpatient treatment
	Mild peri(myo)carditis	Outpatient treatment
	Mild pneumothorax	Outpatient treatment
	Mild pneumonia	Outpatient treatment
	Atrial fibrillation (recurrent/paroxysmal)	Outpatient treatment
	Pericarditis	Outpatient treatment
	Hypertension	Outpatient treatment
	Angina pectoris	Stable, outpatient treatment
	Skeletomuscular	
	Traumatic (mild trauma)	Outpatient treatment
	Gastric/oesophagus problems	
	Mild respiratory problems (such as viral infections)	
	Mental health/panic attack/anxiety disorder	

Diagnosis mentioned above should be linked to the initial complaint of chest pain.

COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; TIA, transient ischaemic attack.

more detailed description of the selection criteria can be found in online supplemental figure S1. The baseline characteristics of the included patients are presented in [table 3](#).

Overall, the median age of the population was 48 years, 56.9% were female, and 39.8% had at least one risk factor for cardiovascular disease. Patients who suffered from an MACE were older, more often male, and more often had a history of cardiovascular disease or predisposing risk factors. Symptom characteristics were also different, with MACE cases more often having heavy/pressure-type chest pain with radiation, nausea and diaphoresis, and less often localised pain that is reproducible with palpation. The mean scores of the clinical risk scores were all higher in the patients who suffered an MACE.

Clinical outcomes

MACE

A MACE was reported in 32 (4.8%) patients within 6 weeks after initial contact. Of those, 6 died (5 from cardiovascular causes), 6 patients had a ST-segment elevation myocardial infarction, 14 non-ST-segment elevation myocardial infarction, 4 unstable angina and 2 cases underwent urgent coronary revascularisation due to crescendo angina.

Other clinical outcomes

At 6 months follow-up, the diagnosis CAD was present in 69 (10.4%) patients. This included 46 (6.9%) patients with an MACE, 10 (1.5%) with cardiac related pain or pressure, 8 (1.2%) with stable angina pectoris and 6 (0.9%) with ischaemic cardiomyopathy.

Table 3 Baseline characteristics of included patients

	Total (n=664)	MACE (n=32)	No-MACE (n=632)
Age, years	48 (32–67)	72 (68–79)	46 (31–66)
Male	286 (43.1%)	20 (62.5%)	266 (42.1%)
Cardiovascular risk factors			
Smoking (current)	119/439 (27.1%)	2/23 (8.7%)	117/416 (28.1%)
Hypertension	169 (25.5%)	18 (56.3%)	151 (23.9%)
Hypercholesterolaemia	83 (12.5%)	8 (25.0%)	75 (11.9%)
Diabetes mellitus	55 (8.3%)	4 (6.3%)	51 (8.1%)
Family history of atherosclerotic disease	54 (8.1%)	2 (6.3%)	52 (8.2%)
Obesity	12 (1.8%)	1 (3.1%)	11 (1.7%)
Atrial fibrillation	63 (9.5%)	7 (21.9%)	56 (8.9%)
History of cardiovascular disease			
Myocardial infarction	43 (6.5%)	6 (18.8%)	37 (5.9%)
CVA/TIA	34 (5.1%)	4 (12.5%)	30 (4.7%)
PAD	10 (1.5%)	1 (3.1%)	9 (1.4%)
PCI	44 (6.6%)	7 (21.9%)	37 (5.9%)
CABG	12 (1.8%)	1 (3.1%)	11 (1.7%)
Use of cardiovascular medications			
Platelet aggregation inhibitor	27 (4.1%)	1 (3.1%)	26 (4.1%)
Salicylates	69 (10.4%)	8 (25.0%)	61 (9.7%)
Statins	119 (17.9%)	9 (28.1%)	110 (17.4%)
Beta-blockers	118 (17.8%)	10 (31.3%)	108 (17.1%)
ACE-inhibitors/ARB	118 (17.8%)	18 (56.3%)	100 (15.8%)
Vitamin K antagonist	48 (7.2%)	4 (12.5%)	44 (7.0%)
NOAC	19 (2.9%)	1 (3.1%)	18 (2.8%)
Nitrates	46 (6.9%)	7 (21.9%)	39 (6.2%)
Chest pain duration			
<1 hour	13 (2.0%)	0 (0.0%)	13 (2.1%)
1–24 hours	317 (47.7%)	16 (50.0%)	301 (47.6%)
>24 hours	240 (36.1%)	14 (43.8%)	226 (35.8%)
Not specified	94 (14.2%)	2 (6.3%)	92 (14.6%)
Chest pain presentation			
Middle or left sided chest pain	302 (45.5%)	14 (43.8%)	288 (45.6%)
Heavy/pressure/tightness	204 (30.7%)	17 (53.1%)	187 (29.6%)
Worse pain with exertion	70 (10.5%)	3 (9.4%)	67 (10.6%)
Pain relieved by nitroglycerin	27 (4.1%)	3 (9.4%)	24 (3.8%)
Radiation of pain to arms/jaw/neck	128 (19.3%)	13 (40.6%)	115 (18.2%)
Nausea or vomiting	85 (12.8%)	6 (18.8%)	79 (12.5%)
Lightheadedness	73 (11.0%)	2 (6.3%)	71 (11.2%)
Diaphoresis	61 (9.2%)	7 (21.9%)	54 (8.5%)
Other relevant symptoms			
Dyspnoea	169 (24.5%)	8 (25.0%)	161 (25.5%)
Cough	80 (12.0%)	0 (0.0%)	80 (12.7%)
Physical examination			
Localised pain	152 (22.9%)	1 (3.1%)	151 (23.9%)
Pain reproducible with palpation	257 (38.7%)	3 (9.4%)	254 (40.2%)
Heart rate, bpm	80 (70–90)	80 (61–95)	80 (70–90)
Systolic blood pressure, mm Hg	136 (120–150)	150 (140–175)	135 (120–150)

Continued

Table 3 Continued

	Total (n=664)	MACE (n=32)	No-MACE (n=632)
Diastolic blood pressure, mm Hg	80 (75–90)	80 (78–100)	80 (75–90)
Pulse oximeter, saturation %	98 (97–99)	97 (95–98)	98 (97–99)
Normal heart sounds	378/389 (97.2%)	18/21 (85.7%)	371/379 (97.9%)
Normal pulmonary sounds	477/527 (90.5%)	20/22 (90.9%)	457/505 (90.5%)
Fever	17 (2.6%)	1 (3.1%)	16 (2.5%)
Scores of clinical decision rules			
Marburg Heart Score	1.27 (1-2)	2.25 (2-3)	1.22 (0.25–2)
INTERCHEST	0.69 (0–1)	2.47 (2-3)	0.60 (0–1)
Gencer rule	2.40 (1-3)	3.72 (2.25–5)	2.33 (1-3)
Bruins Slot rule	4.39 (0–7)	7.94 (5–12.25)	4.21 (0–7)

All categorical data are presented as number and percentage. Continuous variables are presented as median and corresponding 25th and 75th percentile.

ARB, angiotensin II receptor blocker; CABG, coronary artery bypass graft; CVA, cerebrovascular accident; MACE, major adverse cardiac event; NOAC, novel oral anticoagulants; PAD, peripheral artery disease; PCI, percutaneous coronary intervention; TIA, transient ischaemic attack.

A major event within 6 weeks after initial contact was reported in 62 (9.3%) patients. Apart from MACE or CAD, this included 6 (0.9%) cases of atrial fibrillation, 5 (0.8%) patients with pulmonary embolisms and 4 (0.6%) patients with a severe pneumonia.

Common other diagnoses included: chest wall symptoms (46.1%), hyperventilation (4.2%), pneumonia (3.3%), upper gastrointestinal problems (3.3%) and palpitations (3.2%). A complete list of working diagnoses and final diagnoses can be found as online supplemental table S1.

Evaluation of clinical risk scores

The discriminatory properties of the clinical risk scores for 6-week MACE as well as CAD and major events are displayed in figure 1. Overall, only the INTERCHEST score had a point-estimate C-statistic of >0.80, whereas the remaining decision rules performed modestly (0.70–0.80).

Diagnostic performance of all clinical risk scores at various cut-off values (MHS \geq 1, \geq 2 and \geq 3, INTERCHEST score \geq 1 and \geq 2, Gencer rule \geq 2 and \geq 3 and Bruins Slot \geq 2 and \geq 5), as well as corresponding contingency tables are shown in tables 4–6 for MACE, CAD and major events, respectively.

Evaluation of unaided clinical judgement

GPs referred 157 (23.6%) patients for evaluation at the ED. Of those, 26 had an MACE, whereas 131 did not. Of the 507 non-referred patients, 6 turned out to have an MACE within 6 weeks. Based on these data, the sensitivity and specificity of unaided clinical judgement were 81.3% (63.6–92.8) and 79.3% (75.9–82.9), respectively. The corresponding PPV was 16.6% (13.7–19.9) and NPV 98.8% (97.6–99.4).

Clinical risk scores versus physician assessment for MACE

Compared with physician assessment, the MHS using a referral threshold of \geq 1 resulted in fewer missed MACEs

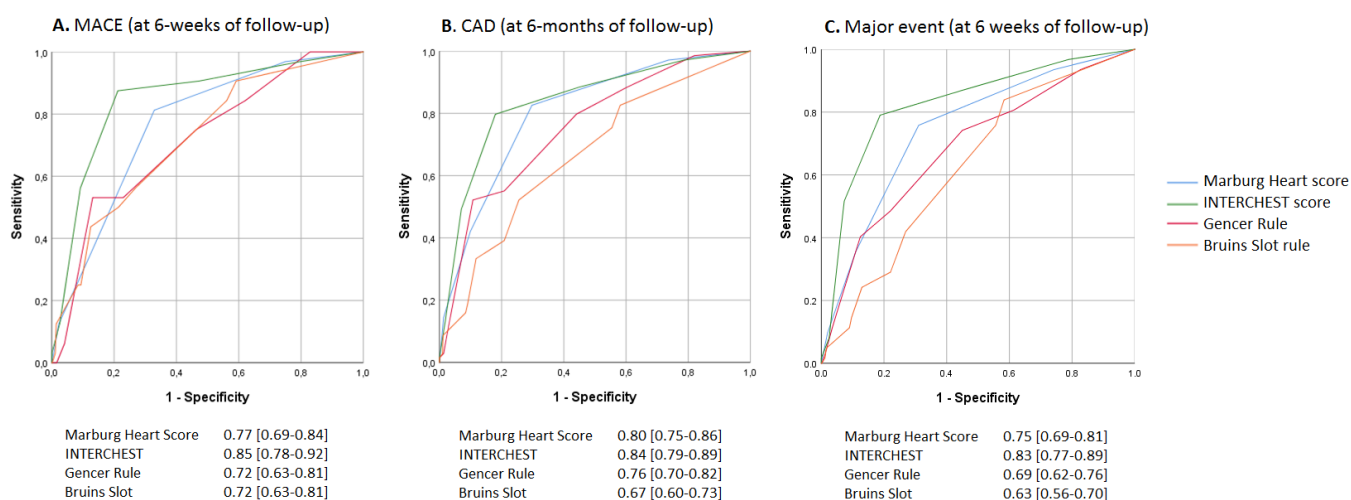


Figure 1 ROC curves and C-statistics of clinical risk scores for: (A) MACE (6 weeks), (B) CAD (6 months) and (C) major event (6 weeks). CAD, coronary artery disease; MACE, major adverse cardiac events.

Table 4 Contingency tables of clinical risk scores at different thresholds for the occurrence of MACE within 6 weeks after presentation

	Threshold	N	MACE	No-MACE	SENS	SPEC	PPV	NPV
Marburg heart score	≥1	505	31	474	96.9%	25.0%	6.1%	99.4%
	0	159	1	158	(83.8–99.9)	(21.7–28.6)	(5.7–6.6)	(95.8–99.9)
	≥2	234	26	208	81.3%	67.1%	11.1%	98.6%
	0–1	430	6	424	(63.6–92.8)	(63.3–70.7)	(9.3–12.3)	(97.2–99.3)
	≥3	88	11	77	34.4%	87.8%	12.5%	96.4%
INTERCHEST	0–2	576	21	555	(18.6–53.2)	(85.0–90.3)	(7.8–19.4)	(95.4–97.1)
	≥1	328	29	299	90.6%	52.7%	8.8%	99.2%
	–1–0	336	3	333	(75.0–98.0)	(48.7–56.6)	(7.8–10.0)	(98.0–99.7)
	≥2	162	28	134	87.5%	78.8%	17.3%	99.1%
Gencer rule	–1–1	502	4	498	(71.0–96.5)	(75.4–81.9)	(14.6–20.3)	(97.4–99.7)
	≥2	420	27	393	84.4%	37.8%	6.4%	98.0%
	0–1	244	5	239	(67.2–94.7)	(34.0–41.7)	(95.5–7.5)	(95.5–99.1)
	≥3	317	24	293	75.0%	53.6%	7.6%	97.7%
Bruins slot	0–2	347	8	339	(56.6–88.5)	(49.7–57.6)	(6.2–9.2)	(95.9–98.7)
	≥2	403	29	374	90.6%	40.8%	7.2%	98.9%
	0	261	3	258	(75.0–98.0)	(37.0–44.8)	(6.4–8.1)	(96.7–99.6)
	≥5	382	27	355	84.4%	43.8%	7.1%	98.2%
Bruins slot	0–4	282	5	277	(67.2–94.7)	(39.9–47.8)	(6.1–8.2)	(96.1–99.2)

MACE, major adverse cardiac events; NPV, negative predictive values; PPV, positive predictive values; SENS, sensitivity; SPEC, specificity.

Table 5 Contingency tables of clinical risk scores at different thresholds for the presence of coronary artery disease within 6 months after presentation

	Threshold	N	CAD	No-CAD	SENS	SPEC	PPV	NPV
Marburg heart score	≥1	505	67	438	97.1%	26.4%	13.3%	98.7%
	0	159	2	157	(89.9–99.7)	(22.9–30.1)	(12.6–14.0)	(95.2–99.7)
	≥2	234	57	177	82.6%	70.3%	24.4%	97.2%
	0–1	430	12	418	(71.6–90.7)	(66.4–73.9)	(21.5–27.5)	(95.4–98.3)
	≥3	88	29	59	42.0%	90.1%	33.0%	93.1%
INTERCHEST	0–2	576	40	536	(30.2–54.5)	(87.4–92.4)	(25.4–41.5)	(91.7–94.3)
	≥1	328	61	267	88.4%	55.1%	18.6%	97.6%
	–1–0	336	8	328	(78.4–94.9)	(51.0–59.2)	(16.8–20.5)	(95.5–98.8)
	≥2	162	55	107	79.7%	82.0%	34.0%	97.2%
Gencer rule	–1–1	502	14	488	(68.3–88.4)	(78.7–85.0)	(29.4–38.8)	(95.6–98.2)
	≥2	420	61	359	88.4%	39.7%	14.5%	96.7%
	0–1	244	8	236	(78.4–94.9)	(35.7–43.7)	(13.2–15.9)	(93.9–98.3)
	≥3	317	55	262	79.7%	56.0%	17.4%	96.0%
Bruins slot	0–2	347	14	333	(68.3–88.4)	(51.9–60.0)	(15.3–19.6)	(93.7–97.5)
	≥2	403	57	346	82.6%	41.9%	14.1%	95.4%
	0	261	12	249	(71.6–90.7)	(37.9–45.9)	(12.7–15.8)	(92.5–97.2)
	≥5	382	52	330	75.4%	44.5%	13.6%	94.0%
Bruins slot	0–4	282	17	265	(63.5–85.0)	(40.5–48.6)	(11.9–15.5)	(91.1–96.0)

CAD, coronary artery disease; NPV, negative predictive values; PPV, positive predictive values; SENS, sensitivity; SPEC, specificity.

Table 6 Contingency tables of clinical risk scores at different thresholds for the presence of a major event within 6 weeks after presentation

	Threshold	N	Major	Non-major	SENS	SPEC	PPV	NPV
Marburg heart score	≥1	505	58	447	93.6%	25.8%	11.5%	97.5%
	0	159	4	155	(84.3–98.2)	(22.3–29.4)	(10.7–12.3)	(93.7–99.0)
	≥2	234	47	187	75.8%	68.9%	20.1%	96.5%
	0–1	430	15	415	(63.3–85.8)	(65.1–72.6)	(17.3–23.2)	(94.7–97.7)
	≥3	88	22	66	35.5%	89.0%	25.0%	93.1%
INTERCHEST	0–2	576	40	536	(23.7–48.7)	(86.3–91.4)	(18.2–33.3)	(91.8–94.2)
	≥1	328	54	274	87.1%	54.5%	16.5%	97.6 %
	–1–0	336	8	328	(76.2–94.3)	(50.4–58.5)	(14.8–18.3)	(95.5–98.7)
	≥2	162	49	113	79.0%	81.2%	30.3%	97.4%
Gencer rule	–1–1	502	13	489	(66.8–88.3)	(77.9–84.3)	(26.0–34.9)	(95.9–98.4)
	≥2	420	50	370	80.7%	38.5%	11.9%	95.1%
	0–1	244	12	232	(68.6–89.6)	(34.6–42.6)	(10.5–13.4)	(92.0–97.0)
	≥3	317	46	271	74.2%	55.0%	14.5%	95.4%
Bruins slot	0–2	347	16	331	(61.5–84.5)	(50.9–59.0)	(12.5–16.8)	(93.1–97.0)
	≥2	403	52	351	83.9%	41.7%	12.9%	96.2%
	0	261	10	251	(72.3–92.0)	(37.7–45.8)	(11.5–14.4)	(93.4–97.8)
	≥5	382	47	335	75.8%	44.4%	12.3%	94.7%
	0–4	282	15	267	(63.3–85.8)	(40.3–48.4)	(10.7–14.1)	(91.9–96.5)

NPV, negative predictive values; PPV, positive predictive values; SENS, sensitivity; SPEC, specificity.

(1 vs 6), however, at the expense of referring almost all patients (505 out of 664). When using a referral threshold of ≥2, we found an identical number of missed cases compared with unaided GP assessment (6 vs 6 MACEs), but still with a higher number of referrals (234 vs 157). When evaluating the INTERCHEST score ≥2, we found a (numerically) lower number of missed cases (4 vs 6 MACEs), with a similar amount of referrals (n=162 vs n=157). The Gencer rule as well as the Bruins Slot rule resulted in higher referral rates, and no improvement in terms of missed cases.

Evaluation of clinical risk scores for CAD

When evaluating the performance of clinical risk scores to identify CAD, the MHS and INTERCHEST rules had NPVs of at least 97% (table 5). The INTERCHEST had the highest PPV compared with the other scores, of which INTERCHEST score of ≥2 points proved to be the optimal balance between PPV and NPV (NPV=97.2%, PPV=34.0%). When applying this threshold, a total of 162 (24.4%) patients would have been identified as being at risk of CAD, of which 55 would have proved to have CAD within 6 months after initial contact.

Evaluation of clinical risk scores for major events

We further explored the performance of clinical risk scores for identifying major (both cardiovascular and non-cardiovascular) events within 6 weeks. Even though cardiovascular events were more common, the results of major events were relatively consistent compared with the

evaluation of CAD. Both the MHS and INTERCHEST rules had NPVs of at least 97% (table 6). The INTERCHEST score of ≥2 points, again proved to be the most favourable with a PPV of 30.3% and NPV of 97.4%. When applying this threshold, 49 out of 162 referred patients would have suffered a major event.

DISCUSSION

Chest pain is a common symptom in urgent primary care with a wide range of possible underlying conditions. We evaluated the diagnostic performance of four primary care-based clinical risk scores in terms of differentiating high-risk versus low-risk cases for MACE. We found that the INTERCHEST score had the best discriminatory properties (C-statistics of >0.80), as well as good sensitivity and specificity. Furthermore, we found that when comparing the INTERCHEST score (cut-off of 2 points) to unaided clinical judgement, one could slightly improve safety (picking up 2 of 6 missed MACE cases), with comparable efficacy (similar referral rate). None of the other clinical risk scores improved safety or efficacy over clinical judgement.

Strengths and limitations

To our knowledge, this study represents the first head-to-head evaluation of four clinical decision aids, conducted exclusively in an out-of-hours acute primary care setting. The study involved a consecutive sample of patients and

was conducted in a large-scale urgent primary care centre, involving over a hundred GPs, with a close to completeness of clinical outcomes data, and is therefore likely a representative sample. Moreover, the availability of follow-up information on both cardiac and non-cardiac conditions allowed us to evaluate the robustness of the clinical decision rules for a multitude of clinical outcomes.

Our study also comes with a number of limitations due to its retrospective design based on routine care data. The foremost limitation is that we could not mandate structural follow-up, for instance, of troponin and/or electrocardiography (which is not routinely performed in Dutch general practice in chest pain), and therefore, any recorded events were clinically driven. Although we used an adequate follow-up duration, some form of verification bias could still be in play. Another limitation was that we had to assume that a risk score element was absent when it was not recorded in the electronic patient chart. Also in the case of 'GP's suspicion of a cardiac substrate', we relied on features of unaided clinical judgement (eg, working diagnoses, management action), which may have introduced an indirect form of incorporation bias in the comparison between unaided clinical judgement and the INTERCHEST score. Furthermore, as the number of MACE was limited, we cannot rule out a certain degree of imprecision in regard to the diagnostic performance of the studied risk scores. Finally, a mentionable number of GPs refused to provide follow-up data of their patients because of the 'opt-out-plus' design of the study, or expressed liability concerns due to the recent implementation of the European privacy regulations.

Prior studies on the MHS

Bösner *et al* developed the MHS for evaluating the likelihood of CAD in office-based primary care.¹¹ The results were promising, with a C-statistic of 0.87, a sensitivity of 87% and specificity of 81% when using a threshold of 3 points. Subsequent studies confirmed the diagnostic performance in office-based, daytime general practice settings.^{12,13} In a systematic review by our group, we found a C-statistic of 0.84–0.90, sensitivity of 87%–91% and specificity of 61%–81%.¹⁴ In the present study, among patients in an urgent primary care setting, the C-statistic for CAD was somewhat lower 0.80 (0.75–0.86), and the optimal threshold was two points (sensitivity 83%, specificity 70%). The MHS was also previously evaluated among patients with possible ACS in general practice. In this study by Schols *et al*, the authors found that the score was an insufficient instrument for safely ruling out ACS, with an NPV of 88% (using a threshold of two points) and did not outperform unaided clinical judgement.¹⁵ Our study, which used MACE as an endpoint, echoes these findings, suggesting that MHS is not a suitable diagnostic instrument in the setting of acute primary care.

Prior studies on the INTERCHEST rule

The INTERCHEST score was developed by Aerts *et al*¹⁶ for CAD risk assessment in 2017 and was validated in the

cohorts of two studies.^{17,18} The authors found a C-statistic of 0.84, sensitivity of 82%–88% and specificity of 74%–82% when using a threshold of 2. The results are similar to the results seen in our study with a C-statistic of 0.85, sensitivity of 88% and specificity of 79% when using a threshold of 2. Two out of five studies providing the data basis for the derivation of the INTERCHEST rule were also used to derive (and validate) the MHS and the Gencer rule.

The INTERCHEST score includes the GP's sense of alarm by administering points when the GP initially suspected a serious condition. A recent study by Barais *et al*¹⁹ evaluated GPs sense of alarm using a validated 'Gut Feelings Questionnaire' in primary care patients presenting with chest pain and dyspnoea. The study showed that the probability of a life-threatening disease increases from 20% to 35% when a sense of alarm is present. The additional value of the GPs gut feeling was also previously underlined in an interview-based study by Hani *et al*.²⁰ In those interviews, GPs emphasised the importance of their prior knowledge of individual patients' risks for CAD. In the light of this prior work, our study showed that the INTERCHEST holds promise as risk assessment tool for MACE in urgent primary care, with an equal referral rate but improved safety.

Prior studies on the Gencer rule

The derivation study of the Gencer rule showed very high discriminatory ability for differentiating CAD from non-CAD, with a C-statistic of 0.95, a sensitivity of 98% and specificity of 71%.²¹ However, in a subsequent validation study, the discriminatory properties were somewhat less impressive (C-statistic of 0.75, sensitivity of 86% and specificity of 47%).²² In this study among urgent primary care patients, we found the Gencer rule to perform similarly with a C-statistic of 0.76 and a sensitivity (80%–88%) and specificity (40%–56%) for CAD. For predicting MACE (instead of CAD), the Gencer rule is not well suited.

Prior studies on the Bruins Slot rule

The Bruins Slot rule is the only prediction rule that was developed specifically to rule out ACS in a primary care setting. The original derivation study showed a C-statistic of 0.66 compared with a C-statistic of the GP of 0.75.²³ In this study, we found that Bruins Slot rule also performed suboptimal for the outcomes MACE, CAD and major events. Considering these studies, the Bruins Slot rule is not preferable for use in urgent primary care.

Implications for clinical practice

GPs hold a low 'failure' rate when it comes to the evaluation of patients with chest pain. A survey among Dutch GPs found that the majority of participants would accept a missed diagnosis rate of 0.1%–1% for ACS, and a maximum of fifty unnecessary referrals for each ACS case.⁵ Furthermore, four out of five GPs would welcome new tools, such as clinical risk scores, to improve decision making.⁵ Overall, the GPs' performance falls within



this range, with 6 missed cases out of 664 cases. Only the INTERCHEST score seems to improve diagnostic accuracy. This may imply further research should focus on the development of a more discriminative clinical risk score to aid GPs in out-of-hours primary care settings.

CONCLUSION

While available risk scores appear safe and have reasonable to good discriminatory properties, they do not outperform unaided clinical judgement for evaluating chest pain in out-of-hours primary care. However, our study suggests that the INTERCHEST score may slightly improve risk stratification as it resulted in numerically fewer missed MACEs. This warrants further prospective evaluation.

Twitter Remco Rietveld @Remco Rietveld

Contributors WAML and REH designed and directed the project and act as guarantors. RR contributed to the implementation of the project and provided input on the cooperation with the out-of-hours primary care centre and affiliated GPs. REH was in charge of overall direction and planning. MK, IS and AM aided in the data collection. MK and AM wrote the manuscript with input from all authors.

Funding Funding for the TRACE study was provided by a grant from the Amsterdam Cardiovascular Sciences Research Institute (ACS-2018-TRACE-22254), a grant from ZonMw-HGOG (project number: 839150004), and internal funding from the department of general practice of the Amsterdam University Medical Centres, location AMC, the Netherlands.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The TRACE study protocol was evaluated by the Medical Ethical Review Committee of the Amsterdam University Medical Centres, location AMC and was exempted full evaluation given the observational nature of the study. We reached out to data protection specialists and legal support how to best handle data protection and privacy of the individuals in the study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Our data consist of deidentified participant data obtained from the out-of-hours primary care facility in Alkmaar, the Netherlands (HONK organisation). Data are available on reasonable request by contacting the corresponding author.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Amy Manten <http://orcid.org/0000-0002-3665-9603>

REFERENCES

- 1 Frese T, Mahlmeister J, Heitzer M, *et al*. Chest pain in general practice: frequency, management, and results of encounter. *J Family Med Prim Care* 2016;5:61–6.
- 2 Hoorweg BB, Willemsen RT, Cleef LE, *et al*. Frequency of chest pain in primary care, diagnostic tests performed and final diagnoses. *Heart* 2017;103:1727–32.
- 3 Haasenritter J, Biroga T, Keunecke C, *et al*. Causes of chest pain in primary care—a systematic review and meta-analysis. *Croat Med J* 2015;56:422–30.
- 4 Biesemans L, Cleef LE, Willemsen RTA, *et al*. Managing chest pain patients in general practice: an Interview-Based study. *BMC Fam Pract* 2018;19:80.
- 5 Harskamp R, van Peet P, Bont J, *et al*. The conundrum of acute chest pain in general practice: a nationwide survey in the Netherlands. *BJGP Open* 2018;2:bjgpopen18X101619.
- 6 Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Stard 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527.
- 7 Manten A, Cuijpers CJJ, Rietveld R, *et al*. Rationale and design of a cohort study evaluating triage of acute chest pain in out-of-hours primary care in the Netherlands (trace). *Prim Health Care Res Dev* 2020;21:e10.
- 8 Castor EDC. Castor electronic data capture, 2019. Available: <https://castoredc.com>
- 9 Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the heart score. *Neth Heart J* 2008;16:191–6.
- 10 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;168:2153–8.
- 11 Bösner S, Haasenritter J, Becker A, *et al*. Ruling out coronary artery disease in primary care: development and validation of a simple prediction rule. *CMAJ* 2010;182:1295–300.
- 12 Haasenritter J, Donner-Banzhoff N, Bösner S. Chest pain for coronary heart disease in general practice: clinical judgement and a clinical decision rule. *Br J Gen Pract* 2015;65:e748–53.
- 13 Haasenritter J, Bösner S, Vaucher P, *et al*. Ruling out coronary heart disease in primary care: external validation of a clinical prediction rule. *Br J Gen Pract* 2012;62:e415–21.
- 14 Harskamp RE, Laeven SC, Himmelreich JC, *et al*. Chest pain in general practice: a systematic review of prediction rules. *BMJ Open* 2019;9:e027081.
- 15 Schols AMR, Willemsen RTA, Bonten TN, *et al*. A nationwide Flash-Mob study for suspected acute coronary syndrome. *Ann Fam Med* 2019;17:296–303.
- 16 International Working Group on Chest Pain in Primary Care (INTERCHEST), Aerts M, Minalu G, *et al*. Pooled individual patient data from five countries were used to derive a clinical prediction rule for coronary artery disease in primary care. *J Clin Epidemiol* 2017;81:120–8.
- 17 Verdon F, Herzig L, Burnand B, *et al*. Chest pain in daily practice: occurrence, causes and management. *Swiss Med Wkly* 2008;138:340–7.
- 18 Bösner S, Becker A, Abu Hani M, *et al*. Accuracy of symptoms and signs for coronary heart disease assessed in primary care. *Br J Gen Pract* 2010;60:e246–57.
- 19 Barais M, Fossard E, Dany A, *et al*. Accuracy of the general practitioner's sense of alarm when confronted with dyspnoea and/or chest pain: a prospective observational study. *BMJ Open* 2020;10:e034348.
- 20 Hani MA, Keller H, Vandenesch J, *et al*. Different from what the textbooks say: how GPs diagnose coronary heart disease. *Fam Pract* 2007;24:622–7.
- 21 Gençer B, Vaucher P, Herzig L, *et al*. Ruling out coronary heart disease in primary care patients with chest pain: a clinical prediction score. *BMC Med* 2010;8:9.
- 22 Bösner S, Becker A, Haasenritter J, *et al*. Chest pain in primary care: epidemiology and pre-work-up probabilities. *Eur J Gen Pract* 2009;15:141–6.
- 23 Bruins Slot MHE, Rutten FH, van der Heijden GJMG, *et al*. Diagnosing acute coronary syndrome in primary care: comparison of the physicians' risk estimation and a clinical decision rule. *Fam Pract* 2011;28:323–8.