Sex-specific evaluation and redevelopment of the GRACE score in non-ST-segment elevation acute coronary syndromes in populations from the UK and Switzerland: a multinational analysis with external cohort validation

Florian A Wenzl, Simon Kraler*, Gareth Ambler, Clive Weston, Sereina A Herzog, Lorenz Räber, Olivier Muller, Giovanni G Camici, Marco Roffi, Hans Rickli, Keith A A Fox, Mark de Belder, Dragana Radovanovic†, John Deanfield†, Thomas F Lüscher†*

Lancet 2022; 400: 744–56

Published Online August 29, 2022 https://doi.org/10.1016/S0140-6736(22)01483-0

See Comment page 710

*Contributed equally

†Jointly directed the study

Center for Molecular Cardiology, University of Zürich, Schlieren, Switzerland (F A Wenzl MD, S Kraler MD, G G Camici PhD, Prof T F Lüscher MD); **Department of Research and Education, University Hospital Zurich, Zurich, Switzerland** (G G Camici); **Department of Statistical Science** (G Ambler PhD) **and Institute of Cardiovascular Sciences** (Prof J Deanfield MD), **University College London, London, UK; Glangwili General Hospital, Carmarthen, UK** (C Weston MD); **Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria** (S A Herzog PhD); **Department of Cardiology, Bern University Hospital, Inselspital, University of Bern, Bern, Switzerland** (L Räber MD); **Department of Cardiology, University Hospital of Lausanne, University of Lausanne, Lausanne, Switzerland** (Prof O Muller MD); **Department of Cardiology, University Hospital Geneva, Geneva, Switzerland** (Prof M Roffi MD); **Cardiology Division, Kantonsspital St Gallen, St Gallen, Switzerland** (H Rickli MD); **Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK** (Prof K A A Fox MB ChB); **National Institute for Cardiovascular Outcomes Research, Barts Health NHS Trust, London, UK** (M de Belder MD); **AMIS Plus**

Summary

Background The Global Registry of Acute Coronary Events (GRACE) 2.0 score was developed and validated in predominantly male patient populations. We aimed to assess its sex-specific performance in non-ST-segment elevation acute coronary syndromes (NSTE-ACS) and to develop an improved score (GRACE 3.0) that accounts for sex differences in disease characteristics.

Methods We evaluated the GRACE 2.0 score in 420 781 consecutive patients with NSTE-ACS in contemporary nationwide cohorts from the UK and Switzerland. Machine learning models to predict in-hospital mortality were informed by the GRACE variables and developed in sex-disaggregated data from 386 591 patients from England, Wales, and Northern Ireland (split into a training cohort of 309 083 [80·0%] patients and a validation cohort of 77 508 [20·0%] patients). External validation of the GRACE 3.0 score was done in 20 727 patients from Switzerland.

Findings Between Jan 1, 2005, and Aug 27, 2020, 400 054 patients with NSTE-ACS in the UK and 20 727 patients with NSTE-ACS in Switzerland were included in the study. Discrimination of in-hospital death by the GRACE 2.0 score was good in male patients (area under the receiver operating characteristic curve [AUC] 0·86, 95% CI 0·86–0·86) and notably lower in female patients (0·82, 95% CI 0·81–0·82; p<0·0001). The GRACE 2.0 score underestimated in-hospital mortality risk in female patients, favouring their incorrect stratification to the low-to-intermediate risk group, for which the score does not indicate early invasive treatment. Accounting for sex differences, GRACE 3.0 showed superior discrimination and good calibration with an AUC of 0·91 (95% CI 0·89–0·92) in male patients and 0·87 (95% CI 0·84–0·89) in female patients in an external cohort validation. GRACE 3·0 led to a clinically relevant reclassification of female patients to the high-risk group.

Interpretation The GRACE 2.0 score has limited discriminatory performance and underestimates in-hospital mortality in female patients with NSTE-ACS. The GRACE 3.0 score performs better in men and women and reduces sex inequalities in risk stratification.

Funding Swiss National Science Foundation, Swiss Heart Foundation, Lindenhof Foundation, Foundation for Cardiovascular Research, and Theodor-Ida-Herzog-Egli Foundation.

Introduction

Non-ST-segment elevation acute coronary syndromes (NSTE-ACS) account for about three quarters of acute coronary syndrome cases in women. The Global Registry of Acute Coronary Events (GRACE) score estimates mortality risk from widely available clinical, electrocardiographic, and biochemical variables and provides an established tool for personalised patient management. Based on evidence from clinical trials,[1–6] selection of invasive treatment strategy, tailored monitoring during hospital stay, and assessment of prognosis according to the GRACE score are recommended across international NSTE-ACS guidelines (class 1a recommendation for guiding treatment).[7,8] Even though the GRACE score was derived from and validated in predominantly male patient populations, it is used in both sexes alike, without accounting for sex-specific disease characteristics of NSTE-ACS.[9] Women with NSTE-ACS display different plaque features and have a higher prevalence of plaque erosion as the primary causative mechanism.[10] Moreover, female patients with NSTE-ACS present at an older age, have a higher comorbidity burden, have longer prehospital delays, and show higher unadjusted mortality risk than do male patients.[11–14] Importantly, female patients are less likely to receive early invasive management.[15]

Despite the growing awareness of its differing discriminative performance in specific patient groups,[16,17] including female patients with ST-segment elevation myocardial infarction,[18,19] the effect of sex differences on

Research in context

Evidence before this study

The Global Registry of Acute Coronary Events (GRACE) score was derived from and validated in predominantly male patient populations but is used in both sexes alike. We searched PubMed on May 11, 2022, without language or date restrictions, for publications on the GRACE score and treatment guidelines in non-ST-segment elevation acute coronary syndromes (NSTE-ACS). The following search terms were used: ("GRACE score") and ("guideline", "acute coronary syndrome" and "ST" and "elevation"). We did not exclude any articles from the search results. Current guidelines for NSTE-ACS recommend basing the selection of the invasive treatment strategy, patient monitoring during hospital stay, and the assessment of prognosis on GRACE risk estimates without accounting for sex (class 1a recommendation for guiding treatment). Baseline features informing GRACE risk estimates differ markedly between sexes and show a sex-specific association with adverse outcomes in patients with acute coronary syndrome. Previous studies report diminished discriminating ability of the GRACE score in female patients with ST-segment elevation myocardial infarction, bringing into question its adequate performance in female patients with NSTE-ACS. The application of machine learning algorithms to medical problems holds promise to improve the prediction of mortality risks across the heterogeneous spectrum of patients with NSTE-ACS.

Added value of this study

In this largest study on the GRACE score to date, including 420 781 female and male patients from England, Wales, Northern Ireland, and Switzerland, we evaluated the performance of the GRACE 2.0 score, characterised baseline risk profiles in a sex-disaggregated manner, and developed a machine learning-based risk score for in-hospital mortality, which captures potential non-linear effects of baseline variables in female and male patients (termed the GRACE 3.0 score). Our study reveals limited discriminatory performance of GRACE 2.0 and underestimation of in-hospital mortality risk in female patients, favouring their incorrect stratification into the low-to-intermediate risk group (GRACE risk ≤3%) in which the score indicates to withhold early invasive treatment. The redeveloped, machine learning-based GRACE 3.0 score was trained, tested, and externally validated in prospectively recruited patients with NSTE-ACS who were undergoing current treatment approaches and showed excellent discriminatory properties. The GRACE 3.0 score classified more female and less male patients as high-risk without leading to increased mortality risk in the low-to-intermediate risk group and provides an updated tool for early risk stratification in patients with NSTE-ACS.

Implications of all the available evidence

Comprehensive meta-analyses of trial data have shown that treatment stratification according to baseline risk leads to improved outcomes in patients with NSTE-ACS. Our study identified sex-related differences in risk assessment by the GRACE 2.0 score and its properties in clinical risk stratification. The newly developed machine learning-based GRACE 3.0 score provides improved predictive performance in both sexes and accounts for sex differences in risk stratification to optimise personalised treatment and to overcome structural inequities in the management of patients with NSTE-ACS. Our study results could inform the design of future trials in patients with NSTE-ACS.

Data Center, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland (D Radovanovic MD); **Royal Brompton and Harefield Hospitals, London, UK** (Prof T F Lüscher); **National Heart and Lung Institute, Imperial College, London, UK** (Prof T F Lüscher); **School of Cardiovascular Medicine and Sciences, Kings College London, London, UK** (Prof T F Lüscher)

Correspondence to:
Prof Thomas F Lüscher, Royal Brompton & Harefield Hospitals, London SW3 6NP, UK
cardio@tomluescher.ch

the performance of the GRACE score in patients with NSTE-ACS remains unclear.[14] Emerging evidence on a distinct profile of baseline risk factors in women and men with NSTE-ACS, the unequal strength of association of individual GRACE components with adverse outcomes,[12,20–22] and the hitherto uniform handling in the GRACE score suggest sex differences in score performance, which might promote structural inequities in the treatment of patients with NSTE-ACS.

In this study, we aimed to evaluate the performance of the GRACE 2.0 score in a sex-disaggregated manner and to redevelop the score in recently recruited populations using machine learning-based approaches to account for interindividual heterogeneity and phenotypic differences between female and male patients with NSTE-ACS.

Methods

Study design and participants

We used current data from 420 781 consecutive patients with NSTE-ACS in nationwide acute coronary syndrome cohorts from the UK and Switzerland. In the UK, patient data were retrieved from the Myocardial Ischaemia National Audit Project (MINAP), a prospective national registry of patients with acute coronary syndrome. The MINAP is the largest single health-care system acute coronary syndrome registry worldwide and covers the entire patient pathway from symptom onset to hospital discharge. Among 1 067 439 patients presenting with acute coronary syndrome to any of the participating hospitals in England, Wales, and Northern Ireland (appendix pp 6–8) between Jan 1, 2005, and March 31, 2017, 400 054 patients had a discharge diagnosis of NSTE-ACS, determined according to the consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee;[23] 97 487 of these patients presented to percutaneous coronary intervention (PCI)-capable university hospitals. In Switzerland, patient data were retrieved from the Acute Myocardial Infarction in Switzerland (AMIS) Plus national registry (NCT01305785)[24] and the Special Programme University Medicine Acute Coronary Syndrome (SPUM-ACS) cohort (NCT01000701).[25] In AMIS Plus, a prospective national registry of patients with acute coronary syndrome in Switzerland, 45 797 patients were admitted

See Online for appendix

to any of the participating hospitals (appendix p 9) between Jan 1, 2005, and Aug 27, 2020, of whom 20 727 had a final diagnosis of NSTE-ACS. The prospective SPUM-ACS registry comprises 4787 consecutive patients with acute coronary syndrome admitted to one of the four major university hospitals in Switzerland who underwent coronary angiography between Dec 8, 2009, and Dec 31, 2017, of whom 2239 had a diagnosis of NSTE-ACS. The cohort profile and detailed inclusion and exclusion criteria of each cohort have been reported previously.[24,25]

Follow-up and assessment of outcomes

The primary study outcome was death in hospital. Additional outcomes were death between hospital admission and 6 months after admission, death between hospital admission and 1 year, death between hospital discharge and 6 months after hospital admission, and death between hospital discharge and 1 year after hospital admission. In-hospital mortality of patients in all cohorts was adjudicated by certified clinicians at the time of the event. Mortality at 6 months and at 1 year in MINAP was ascertained by data linkage to the UK Office for National Statistics using individual patient National Health Service (NHS) numbers. Patients enrolled in AMIS Plus who consented to long-term follow-up at hospital discharge had a scheduled follow-up interview by trained study personnel at 1 year after admission for the index event. In SPUM-ACS, all patients had follow-up visits at 1 month (telephone call) and at 1 year (clinical visit). Additionally, in SPUM-ACS, the 1-year mortality endpoint was reviewed by an external endpoint adjudication committee comprising three certified expert cardiologists who were masked to patient baseline characteristics using prespecified adjudication forms.[25,26]

Evaluation

We calculated the GRACE (version 2.0) score for in-hospital death, death at 6 months, and death at 1 year using the following variables at admission: age, heart rate, systolic blood pressure, Killip class, creatinine concentration, cardiac arrest, presence of ST-segment deviation, and presence of troponin elevation.[9] We evaluated the score in pooled data from the UK and Switzerland; exact numbers of patients available in each analytical cohort at each endpoint are shown in the appendix (p 12). To account for distinct management characteristics of patients receiving a higher hospital level of care (ie, they were treated in a PCI-capable university hospital),[14] we assessed the score performance in all patients and in patients admitted to PCI-capable university hospitals (appendix pp 6–9, 12). The discriminatory performance of the GRACE score was assessed for female and male patients by the area under the receiver operating characteristic (ROC) curve (AUC) and compared using the DeLong test for unpaired ROC curves. Calibration was evaluated graphically using calibration plots and locally weighted scatterplot smoothing-estimated calibration curves. Moreover, a range of performance metrics including the Brier score (average prediction error),[27] accuracy, false omission rate, and the expected–observed ratio were calculated, as reported previously.[28–30] Additionally, we compared mortality risks between female and male patients in predefined clinically meaningful GRACE risk categories for in-hospital deaths (ie, low-to-intermediate risk [≤3%; ≤140 points] and high risk [>3%; >140 points]).[7] Given its broad use in clinical trials and endorsement by treatment guidelines,[1–8] our analyses were primarily aimed at the in-hospital death endpoint of the score. Additional analyses were done for mortality endpoints at 6 months and at 1 year in all patients and in hospital survivors (ie, patients who survived the hospital stay) both on pooled data and at a national level.[9] Further exploratory analyses were done on the patient subgroup receiving PCI treatment (appendix pp 13–14).

Model development and validation

We applied a supervised machine learning approach, called ensemble learning, to capture potential non-linear relationships between patient characteristics and mortality. Ensemble learning combines multiple prediction models to generate better predictions than a single model could.[29,31] Specifically, we applied eXtreme Gradient Boosting (XGBoost; version 1.6.0.1[32]), a widely used[29,33] supervised tree-based learning algorithm, to predict in-hospital mortality in patients with NSTE-ACS. Given their high clinical availability and worldwide use, the eight GRACE variables (age, heart rate, systolic blood pressure, Killip class, creatinine concentration, cardiac arrest, presence of ST-segment deviation, and troponin elevation) were used as model features. Since the risk estimated by the GRACE 2.0 score derives from an unbiased global patient population with acute coronary syndrome and can be calculated internally without requiring additional variables, it was integrated as an additional feature, leading to increased model performance. The derivation cohort (MINAP) contained data from 386 591 patients recruited in England, Wales, and Northern Ireland and was randomly split into two datasets using the Mersenne-Twister random number generator implemented in R, comprising a training cohort (309 083 [80%] patients) that was used to train the machine learning models, and an internal validation cohort (77 508 [20%] patients; appendix pp 3, 5) that was used to test the developed models on unseen data and tune their hyperparameters. The external validation cohort (AMIS Plus) included 20 727 patients from Switzerland. Separate models were trained on male and female patient data. Final models (termed the GRACE 3.0 score) were evaluated based on various performance metrics (including AUC) and calibration plots with a focus on the intermediate-to-high-risk patient population, for whom accurate calibration is most important for clinical risk stratification (appendix p 3). To aid interpretability, the Shapley Additive Explanations (SHAP) approach was used to evaluate the effect of each feature on the model output by assigning an importance value (ie,

	UK (2005–17)						Switzerland (2005–20)					
	All patients (MINAP; n=400 054)			University hospitals (MINAP; n=97 487)			All patients (AMIS Plus; n=20 727)			University hospitals (SPUM-ACS; n=2239)		
	Female (n=145 738)	Male (n=254 316)	p value	Female (n=33 302)	Male (n=64 185)	p value	Female (n=5576)	Male (n=15 151)	p value	Female (n=462)	Male (n=1777)	p value
GRACE 2.0 components												
Age, years	76 (66–84)	69 (58–79)	<0·0001	75 (65–83)	68 (57–77)	<0·0001	74 (64–82)	66 (57–76)	<0·0001	71 (61–79)	64 (54–73)	<0·0001
Heart rate, beats per min	80 (69–95)	76 (65–90)	<0·0001	80 (69–95)	75 (65–89)	<0·0001	78 (68–90)	75 (65–88)	<0·0001	75 (65–84)	75 (65–84)	0·52
Systolic blood pressure, mm Hg	142 (123–161)	139 (122–157)	<0·0001	142 (124–162)	139 (122–156)	<0·0001	141 (123–161)	140 (122–159)	<0·0001	136 (121–152)	130 (116–145)	<0·0001
Creatinine, mg/dL	0·9 (0·7–1·2)	1·1 (0·9–1·3)	<0·0001	0·9 (0·7–1·1)	1·0 (0·9–1·3)	<0·0001	0·8 (0·7–1·0)	1·0 (0·8–1·2)	<0·0001	0·7 (0·6–0·9)	0·9 (0·8–1·0)	<0·0001
Cardiac arrest	826/141 695 (0·6%)	2849/247 683 (1·2%)	<0·0001	192/30 965 (0·6%)	822/59 722 (1·4%)	<0·0001	134/5564 (2·4%)	533/15 106 (3·5%)	<0·0001	4/462 (0·9%)	41/1777 (2·3%)	0·049
ST-segment deviation	34 680/140 188 (24·7%)	61 507/244 550 (25·2%)	0·0044	8107/32 318 (25·1%)	15 694/62 220 (25·2%)	0·64	2100/5559 (37·8%)	5136/15 082 (34·1%)	<0·0001	19/426 (4·5%)	71/1627 (4·4%)	0·093
Troponin >99th percentile	128 134/142 184 (90·1%)	223 376/247 944 (90·1%)	0·78	29 480/32 745 (90·0%)	56 864/62 995 (90·3%)	0·24	2790/2864 (97·4%)	7744/7958 (97·3%)	0·76	397/420 (94·5%)	1547/1638 (94·4%)	0·95
Killip class												
I	52 820/71 840 (73·5%)	104 220/129 598 (80·4%)	<0·0001	13 304/17 980 (74·0%)	28 663/35 574 (80·6%)	<0·0001	4323/5461 (79·2%)	12 600/14 726 (85·6%)	<0·0001	378/441 (85·7%)	1577/1733 (91·0%)	0·0012
II	13 464/71 840 (18·7%)	18 226/129 598 (14·1%)	..	3227/17 980 (17·9%)	4778/35 574 (13·4%)	..	778/5461 (14·2%)	1327/14 726 (9·0%)	..	49/441 (11·1%)	112/1733 (6·5%)	..
III	5166/71 840 (7·2%)	6422/129 598 (5·0%)	..	1313/17 980 (7·3%)	1846/35 574 (5·2%)	..	232/5461 (4·2%)	402/14 726 (2·7%)	..	10/441 (2·3%)	31/1733 (1·8%)	..
IV	390/71 840 (0·5%)	730/129 598 (0·6%)	..	136/17 980 (0·8%)	287/35 574 (0·8%)	..	128/5461 (2·3%)	397/14 726 (2·7%)	..	4/441 (0·9%)	13/1733 (0·8%)	..
Cardiometabolic risk factors												
BMI, kg/m^2	26·5 (23·0–30·9)	27·4 (24·6–30·8)	<0·0001	26·8 (23·2–31·2)	27·4 (24·6–30·8)	<0·0001	25·7 (22·9–29·4)	26·9 (24·6–29·7)	<0·0001	26·2 (23·1–29·4)	27·1 (24·7–30·0)	0·0036
Body surface area*, m^2	1·7 (1·6–1·8)	2·0 (1·8–2·1)	<0·0001	1·7 (1·6–1·8)	2·0 (1·8–2·1)	<0·0001	1·7 (1·6–1·8)	2·0 (1·9–2·1)	<0·0001	1·7 (1·6–1·8)	2·0 (1·9–2·1)	<0·0001
Current smoker	22 755/129 505 (17·6%)	54 500/225 821 (24·1%)	<0·0001	5082/28 639 (17·7%)	13 497/54 943 (24·6%)	<0·0001	2663/5498 (48·4%)	3687/14 964 (24·6%)	<0·0001	131/453 (28·9%)	623/1746 (35·7%)	<0·0001
Total cholesterol, mmol/L	4·8 (3·9–5·8)	4·5 (3·6–5·5)	<0·0001	4·8 (3·9–5·8)	4·5 (3·6–5·5)	<0·0001	5·3 (4·4–6·2)	5·0 (4·2–5·9)	<0·0001	4·9 (4·1–5·7)	4·7 (3·9–5·4)	<0·0001
Type 2 diabetes	32 892/136 422 (24·1%)	58 048/233 411 (24·9%)	<0·0001	7978/30 612 (26·1%)	14 912/57 483 (25·9%)	0·70	1412/5338 (26·5%)	3323/14 540 (22·9%)	<0·0001	99/462 (21·4%)	385/1777 (21·7%)	0·91
HbA_{1c}, %	..	..	..	..	..	..	5·7 (5·4–6·1)	5·7 (5·4–6·2)	0·61	5·9 (5·5–6·6)	5·8 (5·5–6·5)	0·43
Medical history												
Dyslipidaemia†	44 138/130 103 (33·9%)	84 244/223 447 (37·7%)	<0·0001	10 511/29 197 (36·0%)	21 499/55 163 (39·0%)	<0·0001	3088/5504 (56·1%)	9262/15 002 (61·7%)	<0·0001	300/462 (64·9%)	1213/1777 (68·3%)	0·17
Hypertension‡	78 582/132 851 (59·2%)	119 790/227 503 (52·7%)	<0·0001	18 280/29 860 (61·2%)	30 393/56 173 (54·1%)	<0·0001	4059/5330 (76·2%)	9877/14 456 (68·3%)	<0·0001	329/462 (71·2%)	1103/1776 (62·1%)	<0·0001
Previous percutaneous coronary intervention	11 603/131 449 (8·8%)	32 221/225 602 (14·3%)	<0·0001	3169/29 584 (10·7%)	8865/55 689 (15·9%)	<0·0001	884/5402 (16·4%)	3398/14 723 (23·1%)	<0·0001	78/462 (16·9%)	370/1776 (20·8%)	0·059
Previous coronary artery bypass grafting	6371/131 675 (4·8%)	23 946/226 133 (10·6%)	<0·0001	1558/29 587 (5·3%)	6131/55 722 (11·0%)	<0·0001	308/5478 (5·6%)	1383/14 902 (9·3%)	<0·0001	26/462 (5·6%)	120/1777 (6·8%)	0·38

(Table 1 continues on next page)

	UK (2005–17)						Switzerland (2005–20)					
	All patients (MINAP; n=400 054)			University hospitals (MINAP; n=97 487)			All patients (AMIS Plus; n=20 727)			University hospitals (SPUM-ACS; n=2239)		
	Female (n=145 738)	Male (n=254 316)	p value	Female (n=33 302)	Male (n=64 185)	p value	Female (n=5576)	Male (n=15 151)	p value	Female (n=462)	Male (n=1777)	p value
(Continued from previous page)												
Family history of coronary artery disease	28 847/106 572 (27·1%)	58 275/188 527 (30·9%)	<0·0001	7675/24 286 (31·6%)	15 983/46 472 (34·4%)	<0·0001	1534/5482 (28·0%)	4145/14 873 (27·9%)	0·87	141/427 (33·0%)	428/1651 (25·9%)	0·0034
Peripheral vascular disease	5573/129 958 (4·3%)	12 426/222 915 (5·6%)	<0·0001	1436/29 174 (4·9%)	3347/54 959 (6·1%)	<0·0001	416/5491 (7·6%)	1000/14 899 (6·7%)	0·031	36/462 (7·8%)	156/1777 (8·8%)	0·50
Cerebrovascular disease	14 551/131 286 (11·1%)	20 751/225 239 (9·2%)	<0·0001	3188/29 313 (10·9%)	4930/55 204 (8·9%)	<0·0001	405/5491 (7·4%)	1008/14 899 (6·8%)	0·13	19/462 (4·1%)	46/1777 (2·6%)	0·082
Depression	..	..	..	..	..	..	611/4969 (12·3%)	875/13 330 (6·6%)	<0·0001	56/367 (15·3%)	110/1242 (8·9%)	<0·0001
Heart failure§	10 916/131 046 (8·3%)	14 735/224 873 (6·6%)	<0·0001	2188/29 229 (7·5%)	3351/55 072 (6·1%)	<0·0001	225/5390 (4·2%)	539/14 516 (3·7%)	0·13	10/462 (2·2%)	33/1777 (1·9%)	0·67
Chronic kidney disease	10 582/131 148 (8·1%)	17 014/224 915 (7·6%)	<0·0001	2378/29 231 (8·1%)	4237/55 009 (7·7%)	0·026	656/5415 (12·1%)	1298/14 720 (8·8%)	<0·0001	..	..	..
Clinical chemistry and haematology												
C-reactive protein, mg/L	..	..	..	..	..	..	5·0 (2·4–14·0)	5·0 (2·0–11·0)	<0·0001	3·3 (1·6–10·4)	3·2 (1·3–8·6)	0·087
N-terminal-pro hormone BNP, ng/L	..	..	..	..	..	..	2084·0 (627·0–6385·0)	1085·0 (307·8–3634·8)	<0·0001	1263·0 (44·2–2656·5)	409·0 (166·0–1218·8)	<0·0001
Troponin, ng/L¶	7000 (355–232 000)	10 400 (400–240 000)	<0·0001	30 500 (500–356 000)	33 040 (610–391 350)	<0·0001	..	..	..	200 (56–600)	233 (66–687)	0·28
Haemoglobin, g/dL	12·8 (11·5–13·9)	14·0 (12·5–15·1)	<0·0001	12·7 (11·5–13·9)	14·0 (12·5–15·0)	<0·0001	13·1 (12·0–14·1)	14·5 (13·3–15·5)	<0·0001	12·5 (11·6–13·4)	13·9 (12·9–14·9)	<0·0001
Estimated glomerular filtration rate, mL/min/1·73 m²‖	62·1 (43·7–81·1)	72·8 (53·4–89·0)	<0·0001	64·8 (45·5–83·5)	74·9 (55·1–90·6)	<0·0001	69·1 (50·0–86·0)	80·8 (62·5–93·6)	<0·0001	82·0 (63·5–94·6)	89·3 (73·6–99·6)	<0·0001
Left ventricular ejection fraction ≥50%	37 550/59 825 (62·8%)	65 226/108 437 (60·2%)	<0·0001	8762/13 818 (63·4%)	16 652/27 786 (59·9%)	<0·0001	2047/3075 (66·6%)	6212/9271 (67·0%)	0·66	221/292 (75·7%)	880/1140 (77·2%)	0·59
Medication at presentation												
Aspirin	66 684/130 718 (51·0%)	124 635/224 018 (55·6%)	<0·0001	16 502/28 252 (58·4%)	33 309/53 358 (62·4%)	<0·0001	2356/5296 (44·5%)	6279/14 157 (44·4%)	0·87	194/367 (52·9%)	676/1243 (54·4%)	0·61
Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker	52 640/126 213 (41·7%)	90 955/215 996 (42·1%)	0·021	11 661/26 991 (43·2%)	21 880/50 775 (43·1%)	0·77	2527/5576 (45·3%)	5939/15 151 (39·2%)	<0·0001	219/366 (59·8%)	705/1236 (57·0%)	0·34
β-blocker	41 832/126 247 (33·1%)	71 755/216 032 (33·2%)	0·63	9429/26 941 (35·0%)	17 866/50 664 (35·3%)	0·46	2117/5388 (39·3%)	4933/14 395 (34·3%)	<0·0001	166/365 (45·5%)	501/1237 (40·5%)	0·090
$P2Y_{12}$ receptor inhibitor	16 046/114 367 (14·0%)	28 406/196 722 (14·4%)	0·0017	4660/25 348 (18·4%)	8904/48 346 (18·4%)	0·91	574/5111 (11·2%)	1649/13 582 (12·1%)	0·087	61/367 (16·6%)	201/1242 (16·2%)	0·84
Statin	60 315/130 994 (46·0%)	113 504/224 245 (50·6%)	<0·0001	13 421/28 653 (46·8%)	27 292/54 071 (50·5%)	<0·0001	1790/5243 (34·1%)	5474/14 095 (38·8%)	<0·0001	154/366 (42·1%)	638/1242 (51·4%)	0·0018

Data are median (IQR) or n/N (%). AMIS=Acute Myocardial Infarction in Switzerland. GRACE=Global Registry of Acute Coronary Events. MINAP=Myocardial Ischaemia National Audit Project. SPUM-ACS=Special Programme University Medicine Acute Coronary Syndrome. *Estimated according to Du Bois and Du Bois.[36] †Defined as elevation in total cholesterol requiring dietary or drug treatment. ‡Defined as already receiving treatment (drug, dietary, or lifestyle) for hypertension or with consistently recorded blood pressure values exceeding 140/90 mm Hg before admission. §Refers to New York Heart Association class greater than II in AMIS Plus. ¶Refers to peak values in the UK. ||Estimated according to Chronic Kidney Disease Epidemiology Collaboration creatinine equation.[37]

Table 1: **Baseline characteristics of patients with non-ST-segment elevation acute coronary syndromes in the UK and Switzerland, stratified by level of care and sex**

	UK (2005–17)						Switzerland (2005–20)					
	All patients (MINAP; n=400 054)			University hospitals (MINAP; n=97 847)			All patients (AMIS Plus; n=20 727)			University hospitals (SPUM-ACS; n=2239)		
	Female (n=145 738)	Male (n=254 316)	p value	Female (n=33 302)	Male (n=64 185)	p value	Female (n=5576)	Male (n=15 151)	p value	Female (n=462)	Male (n=1777)	p value
GRACE 2.0 risk estimate, %*												
Death in hospital	2·5 (1·1–6·1)	1·7 (0·7–4·2)	<0·0001	2·3 (1·0–5·5)	1·6 (0·7–3·9)	<0·0001	2·1 (1·0–4·4)	1·5 (0·7–3·2)	<0·0001	1·4 (0·8–2·4)	1·0 (0·6–2·0)	<0·0001
Death at 6 months	10·0 (5·0–20·0)	6·0 (3·0–14·0)	<0·0001	9·0 (4·0–18·0)	6·0 (3·0–13·0)	<0·0001	8·0 (4·0–15·0)	5·0 (3·0–11·0)	<0·0001	6·0 (3·0–10·0)	4·0 (2·0–7·0)	<0·0001
Death at 1 year	8·9 (3·7–19·7)	5·2 (2·4–13·2)	<0·0001	7·9 (3·5–17·6)	5·0 (2·3–12·2)	<0·0001	6·6 (3·2–13·1)	4·4 (2·3–9·7)	<0·0001	4·6 (2·7–8·7)	3·1 (1·9–6·1)	<0·0001
Management delay												
Onset-to-door, min	226 (112–669)	218 (106–681)	<0·0001	242 (117–737)	235 (110–752)	0·0061	345 (135–965)	330 (120–960)	0·073	369 (151–735)	286 (120–655)	0·025
Door-to-PCI, min	..	..	..	..	..	..	415 (128–1277)	307 (105–1125)	<0·0001	267 (101–640)	245 (93–588)	0·47
Onset-to-PCI, min	..	..	..	..	..	..	1203 (581–2205)	1035 (480–1940)	<0·0001	876 (599–1222)	757 (464–1122)	0·0036
Early invasive strategy (<24 h)												
All	..	..	..	..	..	..	2875/3624 (79·3%)	9309/11 140 (83·6%)	<0·0001	207/212 (97·6%)	785/809 (97·0%)	0·64
Low-to-intermediate risk (according to GRACE 2.0 score)	..	..	..	..	..	..	1998/2454 (81·4%)	7290/8558 (85·2%)	<0·0001	151/155 (97·4%)	666/687 (96·9%)	0·61
High-risk (according to GRACE 2.0 score)	..	..	..	..	..	..	877/1170 (75·0%)	2019/2582 (78·2%)	0·029	56/57 (98·2%)	119/122 (97·5%)	0·73
Type of intervention												
Coronary angiography	70 666/145 738 (48·5%)	157 354/254 316 (61·9%)	<0·0001	19 657/33 302 (59·0%)	44 809/64 185 (69·8%)	<0·0001	3718/5567 (66·8%)	11 697/15 136 (77·3%)	<0·0001	462/462 (100%)	1777/1777 (100%)	..
PCI	30 454/87 993 (34·6%)	76 896/165 558 (46·4%)	<0·0001	10 427/23 925 (43·6%)	26 609/48 674 (54·7%)	<0·0001	3509/5567 (63·0%)	11 180/15 136 (73·9%)	<0·0001	395/411 (96·1%)	1518/1621 (93·6%)	0·058
Coronary artery bypass grafting	2298/87 993 (2·6%)	8249/165 558 (5·0%)	<0·0001	884/23 925 (3·7%)	3066/48 674 (6·3%)	<0·0001	..	..	..	16/411 (3·9%)	103/1621 (6·4%)	0·058
Thrombolysis	223/126 434 (0·2%)	517/216 489 (0·2%)	0·00014	14/28 753 (<0·1%)	48/54 212 (0·1%)	0·046	34/5398 (0·6%)	115/14 580 (0·8%)	0·25	4/462 (0·9%)	9/1770 (0·5%)	0·37
Procedural characteristics												
Duration of PCI, min	..	..	..	..	..	..	..	..		27·0 (17·0–40·5)	29·0 (18·0–46·0)	0·046
Intra-aortic balloon pump	..	..	..	..	..	..	131/5334 (2·5%)	378/14 433 (2·6%)	0·52	2/462 (0·4%)	30/1776 (1·7%)	0·043
Vasopressor use	..	..	..	..	..	..	270/5233 (5·2%)	697/14 161 (4·9%)	0·50	7/462 (1·5%)	14/1775 (0·8%)	0·15
Aspirin	3933/14 883 (26·4%)	6851/24 247 (28·3%)	<0·0001	1040/2288 (45·5%)	1820/3910 (46·5%)	0·40	4981/5369 (92·8%)	13 808/14 514 (95·1%)	<0·0001	363/460 (78·9%)	1395/1760 (79·3%)	0·66
$P2Y_{12}$ receptor inhibitor	111 877/133 009 (84·1%)	195 841/228 027 (85·9%)	<0·0001	24 402/28 938 (84·3%)	46 870/54 666 (85·7%)	<0·0001	3967/5346 (74·2%)	11 924/14 493 (82·3%)	<0·0001	406/462 (87·9%)	1565/1768 (88·5%)	0·70

(Table 2 continues on next page)

SHAP value) to each feature, as

	UK (2005–17)						Switzerland (2005–20)					
	All patients (MINAP; n=400 054)			University hospitals (MINAP; n=97 847)			All patients (AMIS Plus; n=20 727)			University hospitals (SPUM-ACS; n=2239)		
	Female (n=145 738)	Male (n=254 316)	p value	Female (n=33 302)	Male (n=64 185)	p value	Female (n=5576)	Male (n=15 151)	p value	Female (n=462)	Male (n=1777)	p value
(Continued from previous page)												
Glycoprotein IIb/IIIa receptor inhibitor	3050/123 962 (2·5%)	7673/212 135 (3·6%)	<0·0001	1070/26 977 (4·0%)	2859/50 926 (5·6%)	<0·0001	446/5275 (8·5%)	1745/14 256 (12·2%)	<0·0001	58/462 (12·6%)	286/1770 (16·2%)	0·056
Unfractionated heparin	12 165/122 021 (10·0%)	28 168/208 441 (13·5%)	<0·0001	5305/25 618 (20·7%)	12 598/48 071 (26·2%)	<0·0001	3250/5307 (61·2%)	9670/14 362 (67·3%)	<0·0001	440/460 (95·7%)	1693/1768 (95·8%)	0·92
Low-molecular-weight heparin	72 807/123 689 (58·9%)	121 661/211 601 (57·5%)	<0·0001	15 992/26 281 (60·9%)	29 652/49 559 (59·8%)	0·0064	1582/5278 (30·0%)	3939/14 303 (27·5%)	<0·0001	18/462 (3·9%)	67/1770 (3·8%)	0·91
Fondaparinux	46 190/115 070 (40·1%)	80 923/197 516 (41·0%)	<0·0001	7786/25 049 (31·1%)	14 361/47 352 (30·3%)	0·036	195/4221 (4·6%)	527/11 640 (4·5%)	0·81	28/462 (6·1%)	111/1769 (6·3%)	0·87
Duration of hospital stay, days†	6 (3–10)	5 (3–8)	<0·0001	5 (3–9)	4 (3–8)	<0·0001	5 (3–8)	4 (2–6)	<0·0001	4 (2–6)	3 (2–5)	0·0010
Discharge destination												
Rehabilitation or other hospital	28 130/127 973 (22·0%)	58 866/223 874 (26·3%)	<0·0001	2502/28 969 (8·6%)	5253/55 239 (9·5%)	<0·0001	2610/5012 (52·1%)	7344/13 843 (53·1%)	0·24	252/457 (55·1%)	839/1749 (48·0%)	0·0063
Home	99 843/127 973 (78·0%)	165 008/223 874 (73·7%)	<0·0001	26 467/28 969 (91·4%)	49 986/55 239 (90·5%)	<0·0001	2400/5012 (47·9%)	6492/13 843 (46·9%)	0·23	205/457 (44·9%)	910/1749 (52·0%)	0·0063
Discharge medication												
Aspirin	98 000/110 919 (88·4%)	170 684/185 826 (91·9%)	<0·0001	25 220/27 569 (91·5%)	48 900/52 017 (94·0%)	<0·0001	4857/5249 (92·5%)	13 933/14 526 (95·9%)	<0·0001	454/460 (98·7%)	1739/1756 (99·0%)	0·53
Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker	81 392/131 286 (62·0%)	147 479/223 956 (65·9%)	<0·0001	20 678/29 810 (69·4%)	41 626/56 003 (74·3%)	<0·0001	3914/5238 (74·7%)	10 869/14 443 (75·3%)	0·45	389/459 (84·7%)	1492/1755 (85·0%)	0·89
β-blocker	83 956/106 589 (78·8%)	148 690/179 549 (82·8%)	<0·0001	21 435/26 354 (81·3%)	42 601/50 088 (85·1%)	<0·0001	3974/5234 (75·9%)	10 772/14 451 (74·5%)	0·048	366/459 (79·7%)	1348/1754 (76·9%)	0·19
$P2Y_{12}$ receptor inhibitor	70 255/96 343 (72·9%)	119 844/162 154 (73·9%)	<0·0001	18 484/25 054 (73·8%)	35 030/47 568 (73·6%)	0·70	3905/5140 (76·0%)	11 872/14 156 (83·9%)	<0·0001	410/462 (88·7%)	1624/1777 (91·4%)	0·079
Statin	96 372/110 134 (87·5%)	171 257/184 971 (92·6%)	<0·0001	24 491/27 413 (89·3%)	48 541/51 846 (93·6%)	<0·0001	4398/5223 (84·2%)	13 143/14 458 (90·9%)	<0·0001	442/459 (96·3%)	1709/1756 (97·3%)	0·24
Outcomes												
Major bleed in hospital‡	898/134 587 (0·7%)	1314/229 902 (0·6%)	0·00033	272/30 007 (0·9%)	476/56 271 (0·8%)	0·36	37/5576 (0·7%)	77/15 151 (0·5%)	0·68	17/462 (3·7%)	40/1777 (2·3%)	0·082
Death in hospital	8113/140 732 (5·8%)	9693/245 859 (3·9%)	<0·0001	1366/32 792 (4·2%)	1798/63212 (2·8%)	<0·0001	289/5576 (5·2%)	560/15 151 (3·7%)	<0·0001	2/462 (0·4%)	20/1777 (1·1%)	0·29
Death at 6 months§	20 523/145 738 (14·1%)	25 849/254 316 (10·2%)	<0·0001	3785/33 302 (11·4%)	5291/64 185 (8·2%)	<0·0001	37/1550 (2·4%)	68/4289 (1·6%)	0·042	10/460 (2·2%)	46/1769 (2·6%)	0·60
Death at 1 year§	27 035/145 738 (18·6%)	34 706/254 316 (13·6%)	<0·0001	5132/33 302 (15·4%)	7320/64 185 (11·4%)	<0·0001	64/1543 (4·1%)	134/4258 (3·1%)	0·064	17/442 (3·8%)	68/1743 (3·9%)	0·88

Data are median (IQR) or n/N (%). AMIS=Acute Myocardial Infarction in Switzerland. GRACE=Global Registry of Acute Coronary Events.[9] PCI=percutaneous coronary intervention. SPUM-ACS=Special Programme University Medicine Acute Coronary Syndrome. *Based on complete cases. †Based on data retrieved from patients discharged home. ‡Defined as Bleeding Academic Research Consortium ≥3a. §Refers to post-discharge outcomes in AMIS Plus.

Table 2: **GRACE 2.0 risk, management, and outcomes of patients with non-ST-segment elevation acute coronary syndromes in the UK and Switzerland, stratified by level of care and sex**

described previously.[29] This approach is commonly used in game theory to estimate a player's contribution to success.[29] The ability to predict in-hospital death of the GRACE 3.0 score and the GRACE 2.0 score was compared on unseen data (ie, the internal and external validation cohort) using the DeLong test for paired ROC curves. Reclassification was assessed by comparing the risk groups assigned by GRACE 2.0 and GRACE 3.0, with individuals considered reclassified when groups were discordant. A calculator for the GRACE 3.0 score will be available online.

Statistical analysis

Continuous variables are presented as median and IQR. Categorical data are shown as counts and percentages. Normally distributed variables were compared by Students *t* test, non-normally distributed variables by the Mann-Whitney test, and categorical variables by the χ^2 test, Fisher's exact test, or Kruskal-Wallis test, as appropriate. Where data for the calculation of GRACE 2.0 risk estimates were missing, we applied multiple imputation using chained equations (20 imputations) for each cohort separately. We used predictive mean matching, proportional odds models, and logistic regression models using the binary outcome variables (in-hospital death, death at 6 months, and death at 1 year) as predictors (ie, no imputation was used for these variables in the analysis; appendix p 2).[16] Results obtained in imputed datasets were combined using Rubin's rule[34] to derive an overall estimate and confidence interval (appendix p 2). Given that the XGBoost learning algorithm operates on single datasets, we used a single imputed dataset, generated as described, for training, testing, and external validation of the GRACE 3.0 score. Sensitivity analyses using complete cases were done to explore a potential effect of the imputation on the results (appendix pp 17–21, 23). Internal validation of the GRACE 2.0 score evaluation in complete cases was done in each cohort by using 300 bootstrap samples (appendix p 2).[35] Sex differences in mortality and the importance of individual GRACE features for regression-based mortality predictions were assessed in generalised linear models (appendix p 2). Mortality endpoints with event counts below the predefined minimum of 30 were not considered for analysis. The results were reporte according to the framework for transparant reporting of prediction models summarised in the TRIPOD statement and comply with the STROBE statement (appendix pp 39–41). All p values and CIs are two-sided.

Data were analysed in R version 4.1 and IBM SPSS version 27.0.1. A detailed description of the statistical analyses is presented in the appendix (pp 2–4).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

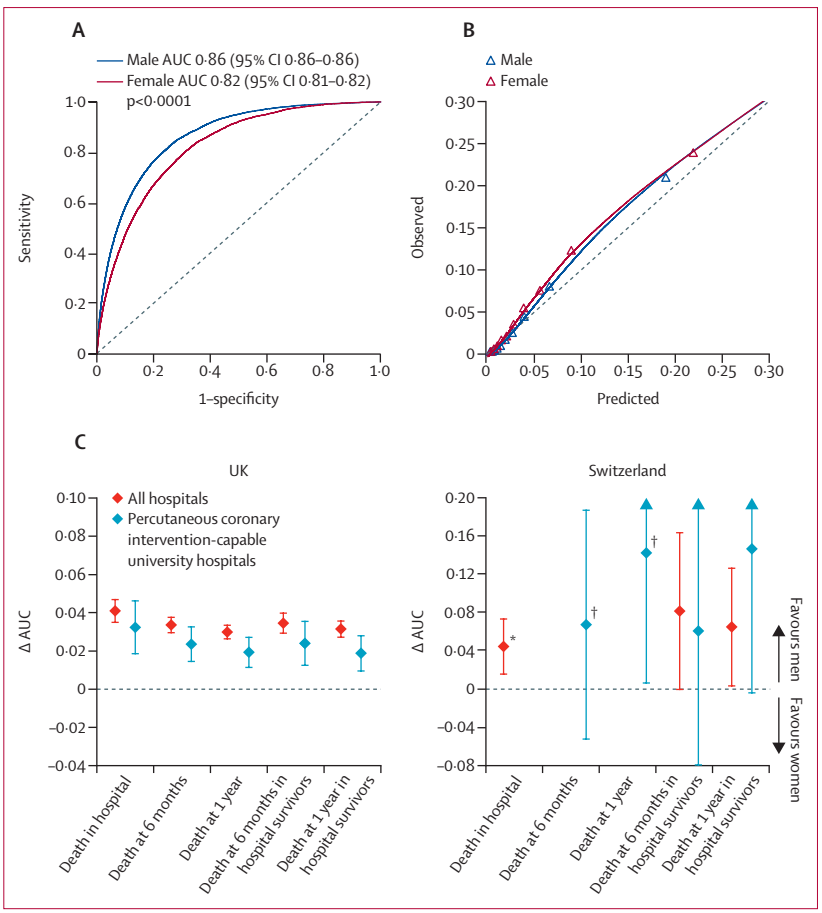

Figure 1: Sex differences in the performance of the GRACE 2.0 score in patients from the UK and Switzerland with non-ST-segment elevation acute coronary syndromes

(A) Receiver operating characteristic curve for the prediction of death in hospital in female and male patients. (B) Observed versus predicted death in hospital. (C) Sex gap in the AUC for the receiver operating characteristic curve of the GRACE model evident at each mortality endpoint. Error bars represent 95% CIs. Δ indicates the difference between male patients and female patients. AMIS=Acute Myocardial Infarction in Switzerland. AUC=area under the curve. GRACE=Global Registry of Acute Coronary Events. *Below event threshold. †Only in-hospital and post-discharge mortality data are available in AMIS Plus.

Results

Between Jan 1, 2005, and Aug 27, 2020, 400 054 patients with NSTE-ACS in the UK and 20 727 patients with NSTE-ACS in Switzerland were included in the study (appendix p 25). At hospital admission, female patients showed marked differences in GRACE components and other baseline risk factors compared with male patients across all cohorts (table 1). Although creatinine concentrations, a component of the GRACE score, suggested better kidney function in female patients compared with male patients, the estimated glomerular filtration rate (eGFR),[37] a sex-adjusted measure of renal function, indicated the opposite. Compared with male patients, the time elapsed between symptom onset and admission was longer and GRACE risk estimates calculated at presentation were consistently higher in female patients with NSTE-ACS (table 2). Paradoxically, female patients were less likely to receive coronary

angiography and to undergo early invasive therapy than were male patients. The female patient population was characterised by longer hospital stays and lower rates of antiplatelet and statin treatment and had a higher crude but not multivariable-adjusted risk for in-hospital mortality compared with male patients (table 2; appendix p 10).

The AUC of the GRACE 2.0 score to predict in-hospital death was 0·86 (95% CI 0·86–0·86) in male patients and 0·82 (95% CI 0·81–0·82; p<0·0001) in female patients. Similar results were obtained in the UK and Switzerland, and for the 6-month and 1-year mortality endpoint across clinical settings of score calculation (figure 1; appendix pp 12, 15–16). Calibration of the GRACE 2.0 score was suboptimal in male and female participants, with higher average prediction errors and false omission rates and lower accuracy in female patients (figure 1; appendix pp 12–16). The GRACE 2.0 score underestimated the in-hospital mortality risk in female patients to a greater extent than in male patients (figure 1), favouring their incorrect stratification to the low-to-intermediate risk group where they were at an increased mortality risk (appendix p 22).

The importance of clinical features informing the GRACE 2.0 model to predict in-hospital death were differentially ranked in regression-based analyses in female and male patients with NSTE-ACS (appendix pp 2, 26), suggesting that sex-specific weighting of GRACE components improves overall model performance. By applying a machine learning algorithm to these features in sex-disaggregated cohorts, we developed and validated the GRACE 3.0 score that, based on an ensemble of decision trees, predicts in-hospital mortality separately in women and men with NSTE-ACS (appendix p 3). The relative feature importance to predict in-hospital death varied across sexes, with marked sex-specific effects of GRACE components on model output (figure 2). For example, baseline heart rate had a non-linear contribution to the model output in both sexes with a higher positive and negative effect on the prediction in male patients with NSTE-ACS.

The GRACE 3.0 score yielded AUCs of 0·89 (95% CI 0·89–0·90) and 0·86 (0·86–0·87) in the training cohort and an AUC of 0·88 (0·87–0·88) and 0·84 (0·83–0·85) in the internal validation cohort in male patients and female patients, respectively (figure 3). When applied to the external validation cohort, the GRACE 3.0 score showed AUCs of 0·91 (95% CI 0·89–0·92) and 0·87 (0·84–0·89) in male patients and female patients, respectively. Discrimination of in-hospital death by the GRACE 3.0 score exceeded that of the GRACE 2.0 score in both validation cohorts irrespective of sex (all p<0·0001). The GRACE 3.0 score showed good calibration (figure 3; appendix p 38), resulting in clinically meaningful differences in the proportion of female and male patients with NSTE-ACS stratified into the high-risk group (figure 4).

Sex-specific GRACE 3.0 risk estimates led to reclassification of women towards the high-risk group and of men towards the low-to-intermediate risk group (figure 4; appendix p 24). As a result, the proportion of patients in the high-risk group increased in female patients and decreased in male patients, without elevating the absolute mortality risk in the low-to-intermediate risk group for either sex.

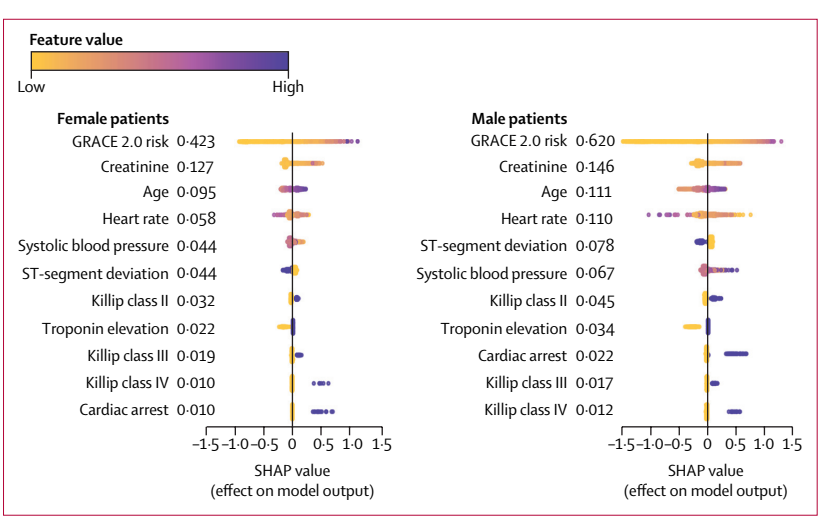

Figure 2: **Feature importance in the GRACE 3.0 score in female and male patients**
The clinical features and the internally calculated GRACE 2.0 risk estimates that inform the supervised machine learning model are ranked according to their contribution to the model output. Each point represents a patient, with colour indicating the feature value. For example, the effect of age on model output is positive when the patient is relatively old (purple) and negative when a patient is relatively young (yellow). Numbers next to the variables represent mean absolute SHAP values. GRACE=Global Registry of Acute Coronary Events. SHAP=Shapley Additive Explanations.

Discussion

In this study, we report results from the largest investigation of the GRACE score to date. We evaluated the GRACE 2.0 score in 420 781 patients with NSTE-ACS from four European countries and found limited discriminatory performance, suboptimal calibration, and underestimation of mortality risk in female patients, who were characterised by a distinct baseline risk profile. Applying a machine learning approach, we derived and externally validated a refined score (termed GRACE 3.0) that appreciates sex-specific relationships between GRACE components and in-hospital mortality, with excellent discriminative ability and good calibration, resulting in improved overall model performance, while relying on identical input variables.

Although sex differences in the clinical characteristics of patients with NSTE-ACS have been reported previously, current data from real-world patient populations in Europe were lacking. In the present study, female patients showed markedly reduced kidney function, evident from lower eGFR.[37] However, creatinine concentrations, a surrogate of renal function used by the GRACE score without accounting for the different physiological range

in women and men,[38] were lower in female patients. Beyond differences in GRACE components, female patients were more likely to present with signs of congestion, had higher N-terminal prohormone of brain natriuretic peptide levels and did not have higher rates of cardiogenic shock (ie, Killip class IV), suggesting that worse Killip class in female patients might be partly related to non-acute coronary syndrome causes.

Female patients also differed in cardiometabolic risk factors, including BMI, cholesterol concentrations, and the prevalence of hypertension, which are not part of the GRACE score. In Switzerland, female patients were almost twice as likely as male patients to present with a history of depression, a patient characteristic with sex-specific associational strength with fatal events after acute coronary syndrome.[22] In the UK, female patients were more likely than male patients to have preserved ejection fraction at baseline, another factor with potentially different prognostic implications after acute coronary syndrome in women and men and not included in the GRACE score. Despite their higher comorbidity burden and increased GRACE 2.0 risk estimates at baseline, female patients were less likely to undergo coronary angiography and to receive early invasive therapy. Although female patients with NSTE-ACS had a higher crude risk for in-hospital mortality relative to male patients, this association was not evident after adjusting for baseline characteristics, in line with observational data of independent cohorts.[11–13]

Beyond the distinct patient risk profiles and management characteristics of female and male patients with NSTE-ACS, the present study unveiled clinically relevant sex-specific limitations of the GRACE 2.0 score. The GRACE 2.0 score showed lower discrimination and suboptimal calibration, as exemplified by a systematic underestimation of in-hospital mortality risk in female patients, thereby expanding on previous reports.[18,19] The dissimilar association between individual GRACE variables and in-hospital mortality in female and male patients with NSTE-ACS in both regression and machine learning models was not considered by the GRACE 2.0 model. This fact might have, at least in part, contributed to diminished performance of the GRACE 2.0 score in female patients with NSTE-ACS, thereby promoting a systematic sex-dependent deviation in early risk stratification and guideline-directed care, probably preventing a subpopulation of female patients with NSTE-ACS from receiving early invasive therapy.

Female patients with NSTE-ACS were less likely to receive early invasive therapy, as reported in previous research.[15,39] Although the management of patients with NSTE-ACS differs from country to country, with a more liberal use of PCI in Sweden and the USA versus the UK,[40] the sex gap in the performance of the GRACE 2.0 score described in this study was evident across geographical boundaries. As diminished GRACE 2.0 score performance in women was independent of geographical region and results obtained in patients undergoing PCI also supported this finding, sex-specific differences in GRACE 2.0 score performance are unlikely to be driven by these factors.

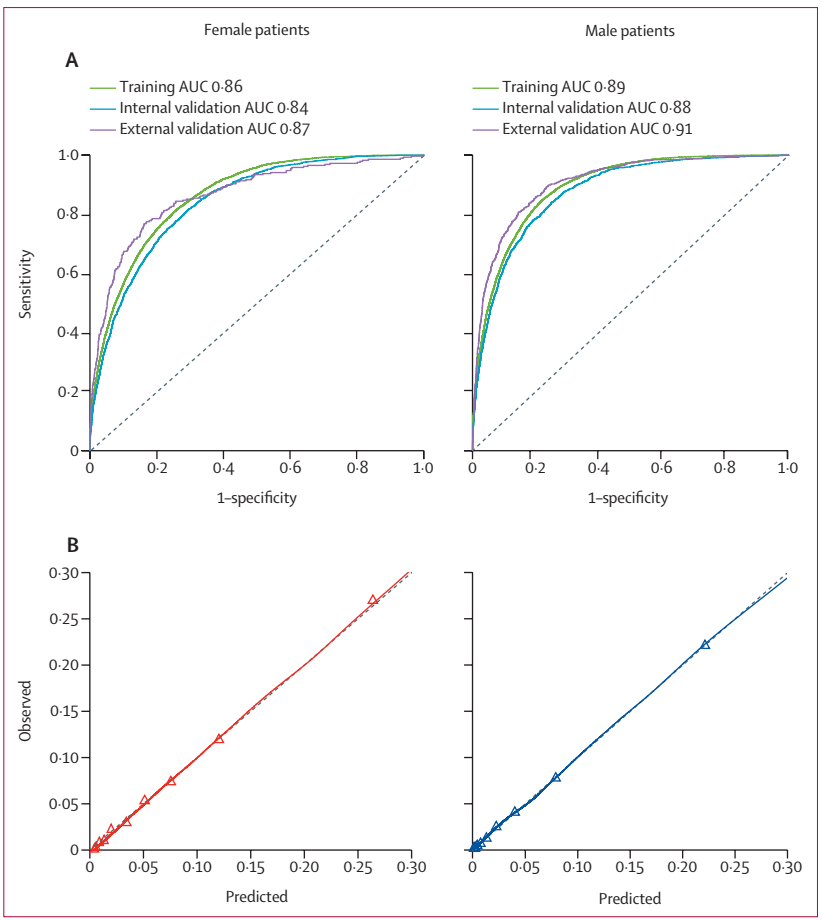

Figure 3: **Performance of the GRACE 3.0 score in female and male patients**
(A) Discrimination of in-hospital death in the training and validation cohorts. (B) Model calibration in unseen data from the internal validation cohort. Triangles indicate predicted and observed probabilities of in-hospital death for patients grouped into tenths using deciles. AUC=area under the curve. GRACE=Global Registry of Acute Coronary Events.

Although the optimal treatment strategy for women with NSTE-ACS remains unclear,[7] our results support the hypothesis that a subpopulation of female patients who were previously classified as low-to-intermediate risk by GRACE 2.0 might benefit from early invasive management. In fact, the updated GRACE 3.0 score stratified more female and less male patients with NSTE-ACS into the high-risk group, with potentially important therapeutic implications.

Although the GRACE score was developed in a predominantly male patient population, recruited until 2007, and is used in both sexes alike, its performance in current female patients with NSTE-ACS was understudied.[9] We delineated sex differences in the importance of score variables and in baseline risk factors beyond the GRACE score, which was reflected in

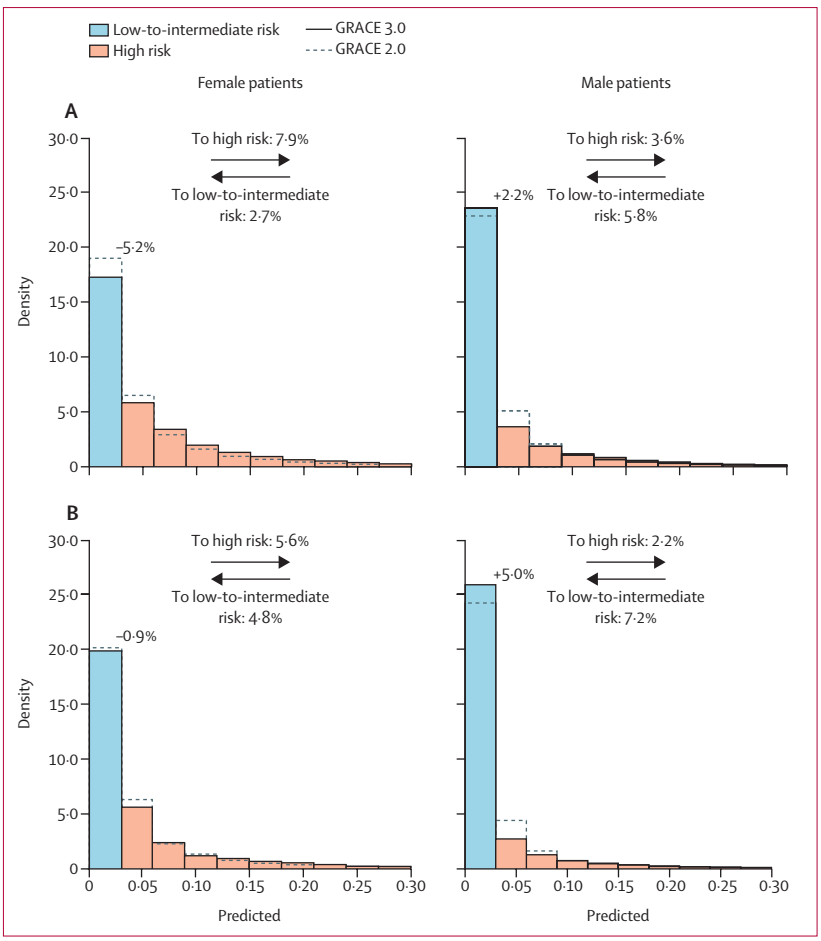

Figure 4: **Effect of the GRACE 3.0 score on risk stratification in female patients and male patients with non-ST-segment elevation acute coronary syndromes**

Histograms showing the distribution of the patient population across 3% intervals of predicted risk in the internal validation cohort from the UK (A) and the external validation cohort from Switzerland (B). Arrows indicate reclassification from the one risk group to another risk group by the GRACE 3.0 score compared with the GRACE 2.0 score.[7,8] Increases and decreases of the patient population in the low-to-intermediate group are indicated by + and –, respectively. Percentages refer to patients of the indicated sex. GRACE=Global Registry of Acute Coronary Events.

unequal score performance. By harnessing machine learning-based methods, we redeveloped the GRACE score, reaching superior performance in internal and external validation datasets, with clinically relevant implications.

Our study has several strengths. First, we analysed the largest patient cohorts in Europe, with a total sample size exceeding previous studies on the GRACE score. Indeed, along with the elegant study by Wilkinson and colleagues,[39] which added an important layer of evidence on the sex gap in guideline-directed care in a cohort containing 418 177 patients with NSTE-ACS, to our knowledge the present study is among the largest investigations into NSTE-ACS. Second, we analysed patients that were enrolled between 2005–20, accounting for the evolution of the NSTE-ACS phenotype and treatment since the end of the recruitment periods of many landmark studies in terms of sex discrepancies in NSTE-ACS and in the derivation and validation cohorts of the GRACE score.[9] Third, we applied resampling techniques to confirm the internal validity of the results. Fourth, we studied the GRACE score at 6 months and at 1 year mortality endpoints in different clinical settings (calculated at presentation and calculated for hospital survivors, respectively), and these findings were largely in line with the results obtained for the in-hospital death endpoint. Consistency of sex differences in the baseline risk profile and in the performance of the GRACE score in NSTE-ACS across independent prospective patient cohorts maximises the external validity of our findings. Despite markedly different mortality rates between the cohorts, probably due to various factors, including differences in management,[40–42] study design,[24,25,43] and unmeasured features of care, consistent underperformance of GRACE 2.0 in female patients with NSTE-ACS was observed. Finally, prospectively collected real-world data, as used in the current study, provides increased generalisability to the European patient population compared with clinical trial data.

There are some limitations inherent to the design of the study cohorts. First, the MINAP and AMIS Plus registries collect data via standardised electronic entry by health-care professionals without complete external event adjudication. Indeed, of the 400 054 patients recruited in MINAP, in-hospital outcomes were only available for 386 591 patients. Additionally, only in-hospital and postdischarge mortality data were available in AMIS Plus, and only a subset of patients were recruited for 1-year follow-up visits. Finally, only data on biological sex but not on the gender of study participants were available, precluding the exploration of socio-cultural influences and transgender people.

In conclusion, the performance of the GRACE 2.0 score is limited by decreased discrimination and underestimation of in-hospital mortality in female patients with NSTE-ACS. The newly developed GRACE 3.0 score accounts for sex-specific weighting of individual GRACE components and shows excellent discrimination and good calibration. Awareness of sex differences in disease biology and the patient risk profile at the time of presentation is critical to improve outcomes in patients with NSTE-ACS. Further external validation is warranted to assess GRACE 3.0 score performance in other populations.

Contributors

FAW, SK, DR, JD, and TFL conceived the study. FAW, SK, SAH, and GA performed data queries, processing, and analyses. FAW and SK wrote the manuscript. FAW and SK have accessed and verified the underlying data reported in the manuscript. All authors vouch for the data and analyses reported. All authors provided important intellectual input in the interpretation of the data, revisited the work critically, approved the final version to be published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the integrity of any part of the work presented are appropriately investigated and resolved. All authors had full access to the data and are responsible for the decision to submit for publication.

Declaration of interests
SK received travel support from the European Atherosclerosis Society and equipment and materials from Roche Diagnostics, outside the submitted work. MR declares institutional research grants from Terumo, Biotronik, Medtronic, Cordis/Cardinal Health, and Boston Scientific, outside the submitted work. LR received funding from Abbott, Biotronik, Boston Scientific, Sanofi, Regeneron, and Heartflow, consulting fees from Abbott, Amgen, AstraZeneca, Canon, NovoNordisk, Medtronic, Sanofi, Occlutech, and Vifor, payment or honoraria from Abbott and Occlutech, and travel support from AstraZeneca. MdB is Chair of the Data Monitoring and Ethics Committee of the UK GRIS Trial and part of the Steering Committee of the DAPA MI Trial. CW is the clinical lead of the MINAP registry. JD received consulting fees from GENinCode UK Ltd, honoraria or consulting fees from Amgen, Boehringer Ingelheim, Merck, Pfizer, Aegerion, Novartis, Sanofi, Takeda, Novo Nordisk, and Bayer, and travel support from the Einstein Professorship Foundation (Berlin, Germany), outside the submitted work. JD holds unpaid leadership positions at Our Future Health and Public Health England. TFL declares institutional educational and research grants from Abbott, Amgen, AstraZeneca, Boehringer Ingelheim, Daichi Sankyo, Novartis, and Vifor, and consulting fees from Daichi Sankyo, Philipps, Pfizer, and Ineeo Inc, outside the submitted work. TFL holds leadership positions at the European Society of Cardiology, Swiss Heart Foundation, and the Foundation for Cardiovascular Research—Zurich Heart House. All other authors declare no competing interests.

Data sharing
Due to data protection regulations related to the different study cohorts involved in this study the authors do not have authorisation to provide unrestricted data access. Requests for the data and additional documents related to the present study should be made to the corresponding author of each single registry (dragana.radovanovic@uzh.ch, j.deanfield@ucl.ac.uk, and cardio@tomluescher.ch).

Acknowledgments
The AMIS Plus registry is funded by unrestricted grants from the Swiss Heart Foundation and from Abbot AG, Amgen AG, AstraZeneca AG, Bayer (Schweiz) AG, Biotronik AG, Boston Scientific AG, B Braun Medical AG, Daiichi-Sankyo/Lilly AG, Cordis Cardinal Health GmbH, Medtronic AG, Novartis Pharma Schweiz AG, Sanofi-Aventis (Schweiz) AG, SIS Medical AG, Terumo AG, and Vascular Medical GmbH, all in Switzerland, and the Swiss Working Group for Interventional Cardiology. The SPUM-ACS registry is primarily supported by the Swiss National Science Foundation (SPUM 33CM30-124112 and 32473B_163271) and the Swiss Heart Foundation (through TFL). Additional support for the SPUM-ACS registry was received from Roche Diagnostics (Boehringer Mannheim; Indianapolis, IN, USA), Eli Lilly (Indianapolis, IN, USA), AstraZeneca (Baar, Switzerland), Medtronic (Münchenbuchsee, Switzerland), Merck Sharpe and Dohme (Lucerne, Switzerland), Sanofi-Aventis (Vernier, Switzerland), and St Jude Medical AG (Zurich, Switzerland). This work was further supported by the Foundation for Cardiovascular Research—Zurich Heart House (through FAW, SK, GGC, and TFL), the Lindenhofstiftung (through TFL and SK), and the Theodor Ida Herzog-Egli Stiftung (through SK). We gratefully acknowledge the invaluable contribution of all health-care professionals who participate in the conduct of the national registries this study is based on. We further express our gratitude towards all administrative personnel and IT personnel involved in data entry and management. Finally, we are grateful to the Epidemiology, Biostatistics and Prevention Institute (University of Zurich, Zurich, Switzerland) for the productive discussions.

References
1 Jobs A, Mehta SR, Montalescot G, et al. Optimal timing of an invasive strategy in patients with non-ST-elevation acute coronary syndrome: a meta-analysis of randomised trials. *Lancet* 2017; **390:** 737–46.
2 Badings EA, The SH, Dambrink JH, et al. Early or late intervention in high-risk non-ST-elevation acute coronary syndromes: results of the ELISA-3 trial. *EuroIntervention* 2013; **9:** 54–61.
3 Thiele H, Rach J, Klein N, et al. Optimal timing of invasive angiography in stable non-ST-elevation myocardial infarction: the Leipzig immediate versus early and late percutaneous coronary intervention trial in NSTEMI (LIPSIA-NSTEMI trial). *Eur Heart J* 2012; **33:** 2035–43.
4 Milosevic A, Vasiljevic-Pokrajcic Z, Milasinovic D, et al. Immediate versus delayed invasive intervention for non-STEMI patients: the RIDDLE-NSTEMI study. *JACC Cardiovasc Interv* 2016; **9:** 541–49.
5 Mehta SR, Granger CB, Boden WE, et al. Early versus delayed invasive intervention in acute coronary syndromes. *N Engl J Med* 2009; **360:** 2165–75.
6 Kofoed KF, Kelbæk H, Hansen PR, et al. Early versus standard care invasive examination and treatment of patients with non-ST-segment elevation acute coronary syndrome. *Circulation* 2018; **138:** 2741–50.
7 Collet J-P, Thiele H, Barbato E, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J* 2021; **42:** 1289–367.
8 Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014; **130:** e344–426.
9 Fox KAA, Fitzgerald G, Puymirat E, et al. Should patients with acute coronary disease be stratified for management according to their risk? Derivation, external validation and outcomes using the updated GRACE risk score. *BMJ Open* 2014; **4:** e004425.
10 Berger JS, Elliott L, Gallup D, et al. Sex differences in mortality following acute coronary syndromes. *JAMA* 2009; **26:** 874–82.
11 Gupta A, Wang Y, Spertus JA, et al. Trends in acute myocardial infarction in young patients and differences by sex and race, 2001 to 2010. *J Am Coll Cardiol* 2014; **64:** 337–45.
12 Champney KP, Frederick PD, Bueno H, et al. The joint contribution of sex, age and type of myocardial infarction on hospital mortality following acute myocardial infarction. *Heart* 2009; **95:** 895–99.
13 Hao Y, Liu J, Liu J, et al. Sex differences in in-hospital management and outcomes of patients with acute coronary syndrome. *Circulation* 2019; **139:** 1776–85.
14 Haider A, Bengs S, Luu J, et al. Sex and gender in cardiovascular medicine: presentation and outcomes of acute coronary syndrome. *Eur Heart J* 2020; **41:** 1328–36.
15 Gupta T, Kolte D, Khera S, et al. Contemporary sex-based differences by age in presenting characteristics, use of an early invasive strategy, and inhospital mortality in patients with non-ST-segment-elevation myocardial infarction in the United States. *Circ Cardiovasc Interv* 2018; **11:** e005735.
16 Hung J, Roos A, Kadesjö E, et al. Performance of the GRACE 2.0 score in patients with type 1 and type 2 myocardial infarction. *Eur Heart J* 2021; **42:** 2552–61.
17 Moledina SM, Kontopantelis E, Wijeysundera HC, et al. Ethnicity-dependent performance of the Global Registry of Acute Coronary Events risk score for prediction of non-ST-segment elevation myocardial infarction in-hospital mortality: nationwide cohort study. *Eur Heart J* 2022; **43:** 2289–99.
18 Gong IY, Goodman SG, Brieger D, et al. GRACE risk score: Sex-based validity of in-hospital mortality prediction in Canadian patients with acute coronary syndrome. *Int J Cardiol* 2017; **244:** 24–29.
19 de-Miguel-Balsa E, Latour-Pérez J, Baeza-Román A, Amorós-Verdú C, Fernández-Lozano JA. GRACE Score validation in predicting hospital mortality: analysis of the role of sex. *J Womens Health (Larchmt)* 2017; **26:** 420–25.
20 Blom MT, Oving I, Berdowski J, van Valkengoed IGM, Bardai A, Tan HL. Women have lower chances than men to be resuscitated and survive out-of-hospital cardiac arrest. *Eur Heart J* 2019; **40:** 3824–34.
21 Mody P, Pandey A, Slutsky AS, et al. Gender-based differences in outcomes among resuscitated patients with out-of-hospital cardiac arrest. *Circulation* 2021; **143:** 641–49.
22 van Loo HM, van den Heuvel ER, Schoevers RA, et al. Sex dependent risk factors for mortality after myocardial infarction: individual patient data meta-analysis. *BMC Med* 2014; **12:** 242.
23 Alpert JS, Thygesen K, Antman E, Bassand JP. Myocardial infarction redefined—a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *J Am Coll Cardiol* 2000; **36:** 959–69.

24 Schoenenberger AW, Radovanovic D, Windecker S, et al. Temporal trends in the treatment and outcomes of elderly patients with acute coronary syndrome. *Eur Heart J* 2016; **37:** 1304–11.

25 Kraler S, Wenzl FA, Georgiopoulos G, et al. Soluble lectin-like oxidized low-density lipoprotein receptor-1 predicts premature death in acute coronary syndromes. *Eur Heart J* 2022; **49:** 1849–60.

26 Laaksonen R, Ekroos K, Sysi-Aho M, et al. Plasma ceramides predict cardiovascular death in patients with stable coronary artery disease and acute coronary syndromes beyond LDL-cholesterol. *Eur Heart J* 2016; **37:** 1967–76.

27 Lee C, Light A, Alaa A, Thurtle D, van der Schaar M, Gnanapragasam VJ. Application of a novel machine learning framework for predicting non-metastatic prostate cancer-specific mortality in men using the Surveillance, Epidemiology, and End Results (SEER) database. *Lancet Digit Health* 2021; **3:** e158–65.

28 D'Ascenzo F, De Filippo O, Gallone G, et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. *Lancet* 2021; **397:** 199–207.

29 Faghri F, Brunn F, Dadu A, et al. Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: a population-based machine-learning study. *Lancet Digit Health* 2022; **4:** e359–69.

30 Katki HA, Kovalchik SA, Petito LC, et al. Implications of nine risk prediction models for selecting ever-smokers for computed tomography lung cancer screening. *Ann Intern Med* 2018; **169:** 10–19.

31 Rokach L. Ensemble-based classifiers. *Artif Intell Rev* 2010; **33:** 1–39.

32 Chen T, Guestrin C. XGBoost. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Aug 13–17, 2016.

33 Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health* 2020; **2:** e516–25.

34 Rubin DB. Inference and missing data. *Biometrika* 1976; **63:** 581–92.

35 Hijazi Z, Oldgren J, Lindbäck J, et al. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. *Lancet* 2016; **387:** 2302–11.

36 Du Bois D, Du Bois EF. A formula to estimate the approximate surface area if height and weight be known. 1916. *Nutrition* 1989; **5:** 303–11.

37 Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; **150:** 604–12.

38 Jones CA, McQuillan GM, Kusek JW, et al. Serum creatinine levels in the US population: third National Health and Nutrition Examination Survey. *Am J Kidney Dis* 1998; **32:** 992–99.

39 Wilkinson C, Bebb O, Dondo TB, et al. Sex differences in quality indicator attainment for myocardial infarction: a nationwide cohort study. *Heart* 2019; **105:** 516–23.

40 McNamara RL, Chung SC, Jernberg T, et al. International comparisons of the management of patients with non-ST segment elevation acute myocardial infarction in the United Kingdom, Sweden, and the United States: The MINAP/NICOR, SWEDEHEART/RIKS-HIA, and ACTION Registry-GWTG/NCDR registries. *Int J Cardiol* 2014; **175:** 240–47.

41 Chung S-C, Gedeborg R, Nicholas O, et al. Acute myocardial infarction: a comparison of short-term survival in national outcome registries in Sweden and the UK. *Lancet* 2014; **383:** 1305–12.

42 Alabas OA, Jernberg T, Pujades-Rodriguez M, et al. Statistics on mortality following acute myocardial infarction in 842 897 Europeans. *Cardiovasc Res* 2020; **116:** 149–57.

43 Herrett E, Smeeth L, Walker L, Weston C. The Myocardial Ischaemia National Audit Project (MINAP). *Heart* 2010; **96:** 1264–67.