BJA

British Journal of Anaesthesia, xxx (xxx): xxx (xxxx)

doi: 10.1016/j.bja.2023.02.030
Advance Access Publication Date: xxx
Clinical Investigation

CLINICAL INVESTIGATION

Risk assessment for major adverse cardiovascular events after noncardiac surgery using self-reported functional capacity: international prospective cohort study

Giovanna A. Lurati Buse[1,*], Eckhard Mauermann[2], Daniela Ionescu[3], Wojciech Szczeklik[4], Stefan De Hert[5], Miodrag Filipovic[6], Beatrice Beck-Schimmer[7], Savino Spadaro[8], Purificación Matute[9], Daniel Bolliger[2], Sanem Cakar Turhan[10], Judith van Waes[11], Filipa Lagarto[12], Kassiani Theodoraki[13], Anil Gupta[14], Hans-Jörg Gillmann[15], Luca Guzzetti[16], Katarzyna Kotfis[17], Hinnerk Wulf[18], Jan Larmann[19], Dan Corneci[20], Frederique Chammartin-Basnet[21], Simon J. Howell[22], and the MET: Reevaluation for Perioperative Cardiac Risk investigators[†]

[1]Anesthesiology Department University Hospital Düsseldorf, Heinrich Heine University, Düsseldorf, Germany, [2]Clinic for Anaesthesia, Intermediate Care, Prehospital Emergency Medicine and Pain Therapy, University Hospital Basel, Basel, Switzerland, [3]Department of Anaesthesia and Intensive Care I, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania, [4]Center for Intensive Care and Perioperative Medicine, Jagiellonian University Medical College, Kraków, Poland, [5]Department of Anaesthesiology and Perioperative Medicine, Ghent University Hospital, Ghent University, Ghent, Belgium, [6]Division of Anesthesiology, Intensive Care, Rescue and Pain Medicine, Kantonsspital St. Gallen, St. Gallen, Switzerland, [7]Institute of Anaesthesiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland, [8]Department of Translational Medicine, University of Ferrara, Ferrara, Italy, [9]Department of Anaesthesia, Hospital Clinic of Barcelona, Universidad de Barcelona, Barcelona, Spain, [10]Department of Anesthesiology and ICU, Ankara University Medical School, Ankara, Turkey, [11]Department of Anesthesiology, University Medical Center Utrecht, Utrecht, the Netherlands, [12]Department of Anesthesiology, Hospital Beatriz Ângelo, Loures, Portugal, [13]Aretaieion University Hospital National and Kapodistrian University of Athens, Athens, Greece, [14]Department of Perioperative Medicine and Intensive Care, Karolinska Hospital and Institution for Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden, [15]Department of Anaesthesiology and Intensive Care Medicine, Hannover Medical School, Hannover, Germany, [16]Anesthesia and Intensive Care Department, University Hospital, Varese, Italy, [17]Department of Anesthesiology, Intensive Therapy and Acute Intoxications, Pomeranian Medical University, Szczecin, Poland, [18]Department of Anesthesiology and Critical Care Medicine, University Hospital Marburg, Marburg, Germany, [19]Department of Anesthesiology, Heidelberg University Hospital, Heidelberg, Germany, [20]Carol Davila University of Medicine and Pharmacy Bucharest Head of Anesthesia and Intensive Care Department I, Central Military Emergency University Hospital "Dr. Carol Davila", Bucharest, Romania, [21]Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, University of Basel, Basel, Switzerland and [22]Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK

*Corresponding author. E-mail: giovanna.luratibuse@med.uni-duesseldorf.de
[†]For the detailed list, please refer to the Supplementary file.

Received: 22 September 2022; **Accepted:** 28 February 2023

Abstract

Background: Guidelines endorse self-reported functional capacity for preoperative cardiovascular assessment, although evidence for its predictive value is inconsistent. We hypothesised that self-reported effort tolerance improves prediction of major adverse cardiovascular events (MACEs) after noncardiac surgery.

Methods: This is an international prospective cohort study (June 2017 to April 2020) in patients undergoing elective noncardiac surgery at elevated cardiovascular risk. Exposures were (i) questionnaire-estimated effort tolerance in metabolic equivalents (METs), (ii) number of floors climbed without resting, (iii) self-perceived cardiopulmonary fitness compared with peers, and (iv) level of regularly performed physical activity. The primary endpoint was in-hospital MACE consisting of cardiovascular mortality, non-fatal cardiac arrest, acute myocardial infarction, stroke, and congestive heart failure requiring transfer to a higher unit of care or resulting in a prolongation of stay on ICU/intermediate care (≥24 h). Mixed-effects logistic regression models were calculated.

Results: In this study, 274 (1.8%) of 15 406 patients experienced MACE. Loss of follow-up was 2%. All self-reported functional capacity measures were independently associated with MACE but did not improve discrimination (area under the curve of receiver operating characteristic [ROC AUC]) over an internal clinical risk model (ROC $AUC_{baseline}$ 0.74 [0.71–0.77], ROC $AUC_{baseline+4METs}$ 0.74 [0.71–0.77], ROC $AUC_{baseline+floors\ climbed}$ 0.75 [0.71–0.78], $AUC_{baseline+fitness\ vs\ peers}$ 0.74 [0.71–0.77], and $AUC_{baseline+physical\ activity}$ 0.75 [0.72–0.78]).

Conclusions: Assessment of self-reported functional capacity expressed in METs or using the other measures assessed here did not improve prognostic accuracy compared with clinical risk factors. Caution is needed in the use of self-reported functional capacity to guide clinical decisions resulting from risk assessment in patients undergoing noncardiac surgery.

Clinical trial registration: NCT03016936.

Keywords: cohort study; effort tolerance; functional capacity; major adverse cardiovascular events; noncardiac surgery; perioperative; postoperative complications; preoperative period; risk assessment

Editor's key points

- Patient-reported functional capacity is an important candidate measure in preoperative risk assessment.
- In this study, patient-reported functional capacity measures were independently associated with major adverse cardiovascular events in that poor self-reported functional capacity influenced risk of adverse events.
- Self-reported functional capacity information did not improve prediction of major adverse cardiac events, i.e. it did not contribute to differentiating patients who will or will not suffer an events over a model of clinical factors.

In Europe, more than 39 million surgeries are performed annually.[1] Preoperative cardiovascular risk assessment informs management and helps patients and physicians to discuss the risk–benefit balance of surgery. Recent European guidelines consider age, cardiovascular disease, cardiovascular risk factors, the procedural risk, and self-reported functional capacity (IIa recommendation) to estimate cardiovascular risk and to guide cardiological work-up.[2]

Whilst the corresponding American guidelines[3] date from 2014 and are centred on functional capacity expressed in metabolic equivalents (METs) (either measured during exercise testing or estimated from patient report), the European guidelines focus on self-reported functional capacity expressed using the ability to climb stairs.[2] The more recent European guidelines refer to the METs study[4] to justify their preference of self-reported measures over exercise testing. Indeed, in that study, whilst independently associated with in-hospital moderate or severe complications, neither peak oxygen consumption nor anaerobic threshold was independently associated with cardiac complication, specifically death or myocardial infarction (MI) and death and myocardial injury at 30 days.[4] Of note, for both approaches (i.e. functional capacity in estimated METs or self-reported functional capacity expressed using the ability to climb stairs), supporting evidence is scarce: most studies assessing self-reported functional capacity for preoperative risk prediction have been comparatively small and showed inconsistent results.[4–8] Whilst associated with 30 day mortality and MI, the Duke Activity Status Index (DASI) did not improve risk classification.[4] A two-centre prospective cohort (258 events; n=4560) in patients aged ≥65 yr or with known cardiovascular disease indicated a significant association between self-reported ability to climb two flights of stairs and suggested a significant improvement in reclassification over the revised cardiac risk index (RCRI).[5] With reference to this study, the European Society of Cardiology guidelines 2022 issued a Class IIa level B recommendation to 'adjust risk assessments according to self-reported ability to climb two flights of stairs … '.[2]

The primary objective of MET: Reevaluation for Perioperative Cardiac Risk (MET-REPAIR) was to assess self-reported measures of effort tolerance[9] for the prediction of major adverse cardiovascular events (MACE) in patients undergoing noncardiac surgery at risk of cardiovascular disease.

Methods

Study design

The MET-REPAIR (NCT03016936) study was an international prospective cohort study in 150 centres (25 countries) between

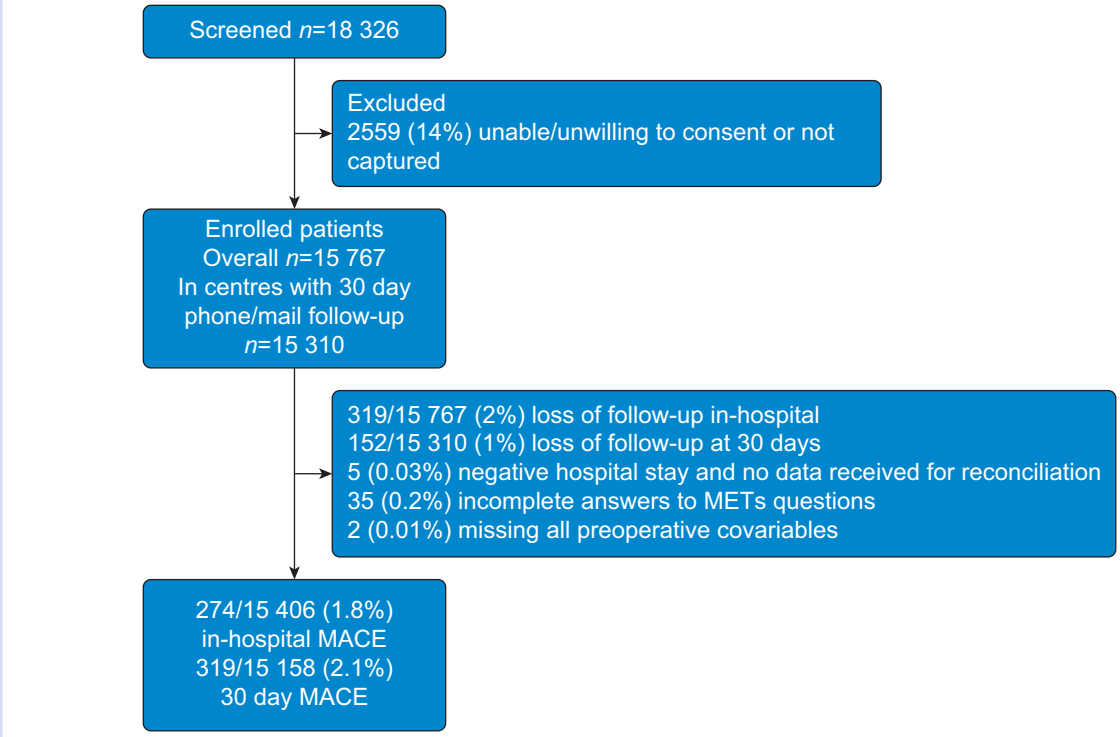

Fig 1. Flowchart. 'Patient screened' refers to the number of patients fulfilling all inclusion criteria and none of the following exclusion criteria (non-elective surgery, acute coronary syndrome or uncontrolled congestive heart failure within 30 days or stroke within 7 days of the planned surgery, no overnight stay, inability to ambulate because of long-standing illnesses or conditions, previous enrolment, and inability to complete the study questionnaire). Centres were asked only to collect separate information on non-participation resulting from inability to give informed consent or unwillingness to participate. MACE, major adverse cardiovascular event; MET, metabolic equivalent.

June 2017 and April 2020. The majority (77.3%) of the centres had referral/tertiary care status.

The study was carried out in accordance with the Declaration of Helsinki and the publicly available research plan (The rotocol and its appendices are avalaiable at the ESAIC CTN homepage). Before enrolment, ethical approval was obtained in all centres. Written informed consent was obtained in 137 centres. In 17 centres, the ethical board stated that written consent was not required.

The project office trained the national principal investigators (PIs) *via* teleconference and provided written definitions and data entry manuals. National PIs provided training to local investigators.

Study population

Patients were identified in the preoperative clinic or from surgical schedules. Patients planned for elective in-patient noncardiac surgery were eligible if aged ≥45 yr and at elevated risk, as defined by either an RCRI ≥2[10] or National Surgical Quality Improvement Program risk calculator for Myocardial Infarction and Cardiac Arrest (NSQIP MICA) >1%,[11] or aged ≥65 yr and undergoing intermediate- or high-risk procedures.[12]

Exclusion criteria were.

(i) Non-elective surgery (within 72 h of diagnosis)
(ii) Acute coronary syndrome or uncontrolled congestive heart failure (CHF) within 30 days, or stroke within 7 days of the planned surgery
(iii) Outpatient surgery (no overnight stay)
(iv) Patients unable to ambulate because of long-standing illnesses or conditions; patients with limited mobility for which they were undergoing surgery were NOT excluded
(v) Inability to complete the study questionnaire
(vi) Inability to give informed consent or unwillingness to participate
(vii) Previous enrolment

Definition and assessment of endpoints

The primary outcome was in-hospital MACE, defined as a composite of intra- or postoperative cardiovascular mortality, non-fatal cardiac arrest, acute MI according to the third universal definition,[13] stroke, and CHF requiring transfer to a higher unit of care or resulting in a prolongation of stay on ICU/intermediate care (≥24 h). Cardiovascular mortality was adjudicated in presence of death after MI,[13] cardiac arrest,

Table 1 Baseline patient characteristics of the entire cohort and stratified by the presence/absence of in-hospital MACE. *Brisk walking, jogging or running, cycling, swimming, or vigorous sports at a comfortable pace or other activities requiring similar levels of exertion. ASA, American Society of Anesthesiologists; eGFR, estimated glomerular filtration rate; MACE, major adverse cardiovascular event (composite of intra- or postoperative in-hospital cardiovascular mortality, non-fatal cardiac arrest, acute myocardial infarction, stroke, and congestive heart failure requiring transfer to a higher unit of care or prolonging stay on ICU/intermediate care ≥24 h); NSQIP MICA, National Surgical Quality Improvement Program risk calculator for Myocardial Infarction and Cardiac Arrest; RCRI, revised cardiac risk index; TIA, transient ischaemic attack.

	Full dataset, *n* (%)	No in-hospital MACE, *n* (%)	In-hospital MACE, *n* (%)
Total	15 406 (100)	15 132 (100)	274 (100)
Female	6059 (39.3)	5969 (39.4)	90 (32.8)
Age (yr)			
40–64	1600 (10.4)	1574 (10.4)	26 (9.5)
65–74	8026 (52.1)	7916 (52.3)	110 (40.2)
≥75	5780 (37.5)	5642 (37.3)	138 (50.4)
ASA physical status ≥3	8882 (57.7)	8682 (57.4)	200 (73.0)
eGFR (ml min^{-1} [1.73 $m]^{-2}$)			
<30/dialysis	768 (5.0)	732 (4.8)	36 (13.1)
30–60	3282 (21.3)	3213 (21.2)	69 (25.2)
≥60	11 254 (73.0)	11 087 (73.3)	167 (61.0)
Diabetes mellitus			
Oral medication	2356 (15.3)	2311 (15.3)	45 (16.4)
Insulin	1337 (8.7)	1297 (8.6)	40 (14.6)
History of coronary heart disease	3700 (24.0)	3593 (23.7)	107 (39.0)
History of congestive heart failure	1920 (12.5)	1847 (12.2)	73 (26.6)
History of peripheral vascular disease	3030 (19.7)	2946 (19.5)	84 (30.7)
History of stroke or TIA	1789 (11.6)	1750 (11.6)	39 (14.2)
Active cancer	7137 (46.3)	6988 (46.2)	149 (54.4)
Chronic obstructive pulmonary disease	2116 (13.7)	2055 (13.6)	61 (22.3)
History of hypertension	11 258 (73.1)	11 045 (73.0)	213 (77.7)
Surgery type			
Low risk	1092 (7.1)	1081 (7.1)	11 (4.0)
Moderate risk	10 086 (65.5)	9953 (65.8)	133 (48.5)
High risk	4228 (27.4)	4098 (27.1)	130 (47.5)
Functional status			
Independent	12 455 (80.8)	12 259 (81.0)	196 (71.5)
Partially independent	2644 (17.2)	2577 (17.0)	67 (24.5)
Fully dependent	307 (2.0)	296 (2.0)	11 (4.0)
Numbers of floors climbed			
>4	3858 (25.1)	3817 (25.3)	41 (15.0)
2–4	9606 (62.4)	9431 (62.4)	175 (63.9)
≤1	1927 (12.5)	1869 (12.4)	58 (21.1)
Self-appraised own cardiorespiratory fitness compared with peers			
Higher	4066 (26.4)	4021 (26.6)	45 (16.5)
Same	6809 (44.2)	6699 (44.3)	110 (40.3)
Lower	4515 (29.3)	4397 (29.1)	118 (43.2)
Self-reported physical activity			
Over 20 min $week^{-1}$*	5201 (33.8)	5149 (34.0)	52 (19.0)
Inactive or low activity	10 197 (66.2)	9976 (66.0)	221 (81.0)
RCRI			
Low (≤1 point)	6212 (40.3)	6142 (40.6)	70 (25.6)
Moderate (2 points)	6075 (39.4)	5976 (39.5)	99 (36.1)
High (≥3 points)	3118 (20.2)	3013 (19.9)	105 (38.3)
NSQIP MICA			
≤1%	7568 (49.1)	7475 (49.4)	93 (33.9)
>1%	7834 (50.9)	7654 (50.6)	180 (65.7)
Unknown	4 (0.0)	3 (0.0)	1 (0.4)

stroke, heart failure or cardiogenic shock, complications of cardiac revascularisation procedure, or death of unknown cause. The definitions for non-fatal cardiac arrest and stroke were in-line with previous outcomes definitions used in noncardiac surgery studies.[14] The same applied to clinical and radiological signs used to define CHF.[14] The decision to add the requirement for a higher level of care for CHF aimed at avoiding the inclusion of less severe congestive events that could easily escape the attention of clinicians, as the study protocol was not mandating radiographic or specific clinical evaluations for cardiac congestion.

Secondary outcomes were MACE at 30 days after surgery and the single items of the composite. Outcomes were adjudicated by the local PI based on standardised definitions (Supplementary material, Section 1.1).

Patients were followed-up in-hospital until discharge, death, or up to 30 days. Of the 150, 148 centres additionally undertook the 30 day follow-up in discharged patients.

Table 2 Adjusted association between various approaches to estimation of functional capacity and in-hospital and 30 day MACE, model discrimination, and model discrimination comparison with the model based on clinical risk factors only for each endpoint. See Supplementary material; adjusted OR for all clinical factors are reported in Supplementary tables. The *P*-values refer to differences in AUC of the models, including various approaches to self-reported functional capacity estimation compared with the AUC of the model based on clinical risk factors only. The baseline model included the following clinical variables: age, ASA physical status class, glomerular filtration rate, diabetes, history of coronary heart disease, history of congestive heart failure, history of peripheral vascular disease, history of stroke or transient ischaemic attack, sex, active cancer, chronic obstructive pulmonary disease, history of hypertension, surgery risk, functional status (independence in activities of daily life), and a random intercept by country. For results of the whole model, please refer to the Supplementary material. Optimized cut-off was calculated by the maximisation of the Youden index. CI, confidence interval; MACE, major adverse cardiovascular event; MET, metabolic equivalent; OR, odds ratio; ROC AUC, area under the curve of receiver operating characteristic.

	In-hospital MACE			30 Day MACE		
	Adjusted OR (95% CI)	**ROC AUC (95% CI)**	**P-value for AUC compared with clinical factors**	**Adjusted OR (95% CI)**	**ROC AUC (95% CI)**	**P-value for AUC compared with clinical factors**
Clinical factors only	See Supplementary material	0.741 (0.710–0.771)	—	See Supplementary material	0.726 (0.698–0.754)	—
Self-reported functional capacity (per MET)	0.92 (0.86–0.97)	0.745 (0.714–0.775)	0.124	0.92 (0.87–0.97)	0.727 (0.699–0.755)	0.368
Self-reported functional capacity (guidelines cut-off in METs)		0.742 (0.712–0.773)	0.321		0.726 (0.698–0.754)	0.470
METs ≥4	1			1		
METs <4	1.61 (1.17–2.23)			1.63 (1.21–2.20)		
Self-reported functional capacity (optimized cut-off in METs)		0.741 (0.711–0.772)	0.291		0.726 (0.698–0.754)	0.527
METs ≥6	1			1		
METs ≤5	1.14 (0.85–1.54)			1.11 (0.84–1.46)		
Number of floors continuously climbed without having to stop		0.746 (0.715–0.776)	0.074		0.730 (0.702–0.758)	0.102
>4	1			1		
2–4	1.39 (0.97–1.99)			1.34 (0.96–1.88)		
≤1	1.90 (1.17–3.11)			1.97 (1.26–3.09)		
Self-appraised cardiorespiratory fitness compared with peers		0.743 (0.714–0.773)	0.236		0.728 (0.701–0.756)	0.264
Higher	1			1		
Same	1.39 (0.98–1.99)			1.47 (1.05–2.04)		
Lower	1.73 (1.17–2.54)			1.71 (1.20–2.45)		
Self-reported regular physical activity		0.753 (0.724–0.782)	0.009		0.731 (0.704–0.759)	0.081
Over 20 min week^{-1}	1			1		
Inactive or low activity	2.02 (1.46–2.79)			1.62 (1.21–2.16)		

Explanatory variables

Self-reported effort tolerance quantified in METs and estimated using a 10-item questionnaire[9] (Supplementary material, Section 1.2) was the main explanatory variable. The MET estimates for single activities from the Compendium of Physical Activities[15] were used to construct 10 questions based on activities needing an intensity of physical effort between one and 10 METs. The questionnaire focused on daily life and household-related activities.[9] With a focus on sensitivity, in the primary analyses, the question corresponding to the highest level of exertion that was answered with 'yes' without any preceding 'no' represented the patient's functional capacity. Sensitivity analyses were conducted, in which we used the highest level of exertion answered with 'yes' independent of preceding 'no'.

The number of floors climbed without having to rest, self-perceived own cardiopulmonary fitness compared with their peers,[16] and weekly physical activity[17] (Supplementary material, Section 1.2.) were also assessed.[9]

Patients completed the questionnaire no more than 30 days before surgery.

Baseline statistical model

The baseline model included the following independent variables: age, sex, ASA physical status class, estimated glomerular filtration rate, active cancer, type of surgery (low-, moderate-, or high-risk procedure),[12] diabetes mellitus, hypertension, CHF, coronary artery disease (CAD), chronic obstructive pulmonary disease, peripheral vascular disease, and stroke (see Supplementary methods, Section 1.3). Data were extracted from medical charts. The selection of covariables and their categorisation were based on the literature.[10,11,18] We decided to use established predictors for MACE after noncardiac surgery as covariables rather than primarily using the RCRI or the NSQIP MICA because of the following considerations: (i) the study endpoint was not congruent with the endpoint for which those scores were developed and validated[10,11] and (ii) the limited discrimination of the RCRI in previous studies.[19–21]

Subgroup analyses

Subgroup analyses were planned for major orthopaedic surgery, major vascular surgery attributable to potential impairment by musculoskeletal discomfort and lower-limb ischaemia, and thoracic surgery. Further, we planned to assess the subgroup of 'cardiovascular healthy' older patients, defined by ≥65 yr with an RCRI <2 and an NSQIP <1%.

Determination of sample size

Assuming an incidence of 2% for the primary composite endpoint,[18,22–24] we aimed to recruit 15 000 patients and observe 300 events to develop a predictive logistic model, including up to 30 predictors to satisfy the widely used rule of 10 events per variable.[25]

Statistical analysis

Baseline characteristics and clinical outcomes were summarised as counts and percentages. The optimal cut-off points for self-reported functional capacity in METs were determined by maximising the Youden index (kernel smoothed densities).[26] A predefined cut-off was 4 METs.[3,12] We used mixed-effects logistic regression to model binary outcomes, where the log for the outcome was modelled as a linear function of a mixture of fixed effects of the independent predictors and a random effect for country variability. Multiple imputation (20 datasets) using chain equations[27] was used to impute missing information on the prognostic factors. Missingness was assumed to be at random, and pre-specified rule for imputation was that missing factors did not exceed 40%.[28]

Covariates were defined and categorised *a priori*.[10,11,18,29] However, we assessed the association of the pre-selected variables in univariable preliminary analyses, and predictors with predefined multiple categories were simplified by combining adjacent categories if their univariable effects were similar (Supplementary methods, Section 1.3). Model performance was primarily assessed using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve in the complete dataset. We tested for AUC superiority using the DeLong test for two correlated ROC curves.[30] Upon suggestion during the review process, decision curves were calculated to provide an additional approach to assess predictive performance. Model internal validation was performed using bootstrap resampling methods at the cluster level (country).[31] The model was internally validated using optimism-corrected concordance index performance (AUC), mean Brier score, and calibration intercept and slope. We combined internal validation and multiple imputation with bootstrap followed by multiple imputation.[32] All statistical analyses were performed using R version 4.0.3 (2020-10-10). The analysis plan is posted on the MET-REPAIR homepage (https://www.esaic.org/research/clinical-trial-network/ongoing-trials/met-repair/study-protocol-and-appendices/).

Sensitivity analyses of the main model

In a sensitivity analysis, we refitted the developed models using the maximum functional capacity reported on the patient questionnaire. Robustness to imputation was assessed by comparing performance of models fitted in the complete dataset and in the imputed dataset.

Justification of deviation from the analysis plan in the protocol

The study protocol included estimation of the net reclassification index after addition of functional capacity to the RCRI and the NSQIP MICA score. In light of controversies regarding this measure,[33–35] we decided *a priori* to not undertake this analysis. *A posteriori* sensitivity analyses were conducted using the RCRI and NSQIP MICA as baseline model. The decision to run the analysis on the single components of the composite only for the outcome 'cardiac death' was dictated by the limited number of the other single events. For the same reason, we decided not to conduct the planned subgroup analyses (orthopaedic surgery, 31 MACE in 2756 patients; vascular surgery, 28 MACE/1019; thoracic surgery, 16 MACE/1050; and ≥65 yr with an RCRI <2 and an NSQIP <1%, 34 MACE/3472).

Results

Descriptive data

The study flowcharts report the number of eligible, included, and analysed patients (Fig 1); of note is the analysis of in-hospital MACE based on the data of all 150 centres (15 406

patients with complete follow-up out of 15 767 enrolled patients). The analysis addressing the secondary endpoint of 30-day MACE, was based on the patients enrolled in the 148 centres, who at the beginning of the study had committed to conducting a 30-day follow-up after discharge as well (15 158 patients with complete 30-day follow-up out of 15 310 enrolled patients). Covariable completeness was above 99%. Baseline characteristics are reported in Table 1 and Supplementary Table 1. There were 315 in-hospital events in 274 of 15 406 patients. These consisted of 123 (0.8%) cardiac deaths, 70 (0.45%) MIs, 41 (0.27%) non-fatal cardiac arrests, 48 (0.31%) cases of heart failure requiring transfer to a higher unit of care or prolonging stay on ICU/intermediate care (≥24 h), and 33 (0.21%) strokes.

Prediction of cardiovascular events using questionnaire-assessed functional capacity in METs

The adjusted odds ratios and 95% confidence intervals (CIs) of functional capacity in METs for in-hospital and 30 day MACEs are shown in Table 2. The ROC AUC of the model, including continuous METs, was 0.745 (95 % CI: 0.714–0.775) for in-hospital MACE (P=0.124 compared with the baseline model). Also, functional capacity dichotomised at 4 METs[3,12] was independently associated with both in-hospital and 30 day MACEs (Table 2). Optimal cut-off according to the maximum Youden indexes was 5 METs. Functional capacity ≤5 METs was associated with in-hospital MACE only in univariable analysis (Supplementary Table 2; Table 2). Detailed results are provided in Supplementary Tables 3–7. Brier score (within cluster, imputed dataset) for the model using self-reported METs dichotomised at 4 for MACE prediction was 0.012, calibration slope 1.011, and intercept 0.008.

The sensitivity analysis using the maximum functional capacity yielded similar findings (Supplementary Fig 1).

The sensitivity analyses using the RCRI and the NSQIP MICA as baseline model confirmed the independent association between continuous and dichotomised at 4 METs when added to the RCRI (+age) and to the NSQIP MICA, respectively (Supplementary Tables 8 and 9). The addition of METs to the RCRI (+age) improved discrimination (P=0.041), but effect size was limited (ROC AUC $0.687_{4MET+RCRI+age}$ [95% CI: 0.654–0.720] *vs* ROC AUC $0.675_{RCRI+age}$ [95% CI: 0.642–0.709]) (Supplementary Table 8). The addition of METs did not significantly increase discrimination over the discrimination of the NSQIP MICA (ROC $AUC_{NSQIP\ MICA}$ 0.672 [95% CI: 0.640–0.705]) (Supplementary Table 9) that in this external cohort was more limited than in the original derivation/validation cohort.[11]

Prediction of cardiovascular events using ability to climb stairs

The ability to climb one or less floor (corresponding to two flights of stairs) without resting was independently associated with in-hospital MACE and 30 day MACE (Table 2); however, stair climbing information did not significantly improve model discrimination when compared with a model based on clinical risk factors alone (Table 2). The sensitivity analyses using the RCRI (+age) and the NSQIP MICA as baseline model indicated a significant independent association between number of stairs climbed and MACE. Stair climbing information improved discrimination when added to RCRI and age (P=0.016); however, effect size was limited (ROC $AUC_{floors+RCRI+age}$ 0.691 [0.660–0.722] *vs* ROC AUC $0.675_{RCRI+age}$ [95% CI: 0.642–0.709]). Stair climbing information did not improve discrimination over the NSQIP MICA (Supplementary Tables 8 and 9).

Prediction of cardiovascular events using other measures of functional capacity

Both self-appraised own cardiorespiratory fitness as being lower than peers and inactivity or performing only limited physical activity regularly (compared with ≥20 min $week^{-1}$ of brisk walking, jogging or running, cycling, swimming, or vigorous sports at a comfortable pace or other activities requiring similar levels of exertion) were independently associated with in-hospital MACE and 30 day MACE (Table 2). Only information on regular physical activity significantly improved model discrimination when compared with the baseline model (Table 2). However, the effect size (delta c-statistics) was limited. Brier score (within cluster, imputed dataset) for the model using regular physical activity dichotomised for MACE prediction was 0.012, calibration slope 1.041, and intercept 0.008. The sensitivity analyses using the RCRI (+age) and the NSQIP MICA as baseline indicated a significant independent association of each alternative self-reported functional capacity measure with MACE. Both alternative measures of functional capacity significantly increased discrimination when added to RCRI (+age); however, discrimination of RCRI-based models remained limited. None of the alternative measures improved discrimination when added to the NSQIP MICA (Supplementary Tables 8 and 9). Decision analysis curves for the various models for in-hospital and 30 day MACE, respectively, are shown in Supplementary Figures 2 and 3. Decision analysis curves did not indicate any relevant benefit of self-reported functional capacity measures over clinical risk factors across misclassification cost ranging up to 30% (decision curves). A misclassification cost of 5% (1/20) refers to the willingness to accept 20 false positives per each true positive; a misclassification cost of 30% (1/3.3) is the acceptance of 3.3 false positives per each true positive.

Discussion

Main findings

The main findings of this analysis are that for patients undergoing elevated risk noncardiac surgery, functional capacity in METs estimated using self-reported tolerance of reference activities was independently associated with in-hospital MACE and MACE at 30 days. However, it did not improve discrimination over a model based on clinical risk factors. Hence, whilst limited self-reported functional capacity increased the risk of MACE, this information did not contribute to the differentiation of patients more likely to experience MACE. The same applied to questionnaire-estimated functional capacity dichotomised at the cut-off of 4 METs, as endorsed by American guidelines.[3] The inability to climb one floor (corresponding to two flights of stairs), as recently endorsed by European guidelines,[2] was independently associated with in-hospital and 30 day MACE. However, this information did not improve discrimination for MACE. Lower self-perceived fitness compared with peers and limited regular physical activity were each independently associated with in-hospital and 30 day MACE. Level of regular physical activity significantly improved the discrimination for in-hospital MACE; however, effect size was limited. In other words, information on the level of regular physical activity statistically

contributed to the differentiation of patients more likely to suffer a MACE; however, improvement was as limited as to probably lack any clinical relevance. All self-reported functional capacity measures improved discrimination over the RCRI, but the effect size was limited. Discrimination over the NSQIP MICA was not improved.

Comparison to previous studies

Studies using estimated METs

In the Measurement of Exercise Tolerance before Surgery (METS) study,[4] physician-estimated METs (i.e. assessed in unstructured interviews using questions at the discretion of the attending physicians) were independently associated with a composite endpoint of neither all-cause mortality and MI nor all-cause mortality and myocardial injury at 30 days. In our cohort, self-reported functional capacity in METs assessed using a structured interview was independently associated with in-hospital MACE and 30 day MACE. This divergence may arise from different approaches used for METs estimation (physician estimated/unstructured *vs* structured), different endpoints, a greater burden of comorbidity in the MET-REPAIR population (e.g. 12% *vs* 24% CAD), and different numbers of events (28/1351 *vs* 274/15 406). Marsman and colleagues[6] presented a secondary analysis within a single-centre cohort of 4879 patients undergoing noncardiac surgery. Whilst independently associated with myocardial injury after noncardiac surgery, postoperative MI, and all-cause death, subjective functional capacity in METs did not improve discrimination compared with the RCRI.[6] Sensitivity analyses within our study suggest that functional capacity expressed in METs improved discrimination compared with the RCRI. However, (i) discrimination of the RCRI-based model was limited; and (ii) in the main analysis using the study baseline model, neither continuous nor dichotomised METs significantly improved model discrimination over a clinical risk factor model.

Stair climbing for estimation of functional capacity

In a secondary analysis of a prospective cohort,[5] the inability to climb two flights of stairs was independently associated with a composite of cardiac death and major adverse cardiac events, and it improved risk classification over the RCRI. Whilst based on a large number of events, the study[5] suffered from the fact that it was conducted in two centres only and that it used the RCRI as baseline for reclassification in spite of its limited discrimination (<0.7)[35] in the study sample.

The present study confirmed the independent association between limited stair climbing ability and MACE after noncardiac surgery. A discrimination improvement was established only over the RCRI (in line with the study mentioned earlier, albeit numerically more limited)[5] but not over the full clinical model or over the NSQIP MICA. Further, discrimination of the RCRI-based models remained limited and delta c-statistics by the addition of stair climbing to RCRI and age very modest.

Other approaches to the estimation of functional capacity

In the METS study, the DASI, a structured interview to assess functional capacity, was independently associated with 30 day mortality and MI, but it did not significantly improve reclassification over the RCRI.[4] Similarly, in our study, there was a significant association between various measures of functional capacity assessed using a structured questionnaire and MACE; however, this information did not significantly improve model discrimination over clinical risk factors in the form of the study baseline model or the NSQIP MICA. In contrast, all self-reported measures of functional capacity improved discrimination compared with the RCRI (+age).

In line with data for long-term cardiac death in the non-operative setting,[16] lower self-perceived own cardiorespiratory fitness compared with peers was independently associated with in-hospital and 30 day MACE in the present study. This underlines patients' self-awareness and does not support concerns of observer bias in relation to self-reported functional capacity.

In the general population, lower physical activity is associated with mid-to long-term all-cause and cardiovascular mortality and cardiovascular disease.[36,37] Whilst a previous cohort suggested a relevant proportion of sedentary lifestyle in patients planned for noncardiac surgery,[38] the authors did not assess its impact on outcome. The contribution of self-reported level of physical activity towards cardiovascular risk prediction in the noncardiac surgery setting needs confirmation in future studies. Given the lack of other setting-specific studies and the limited effect size, external confirmation of our findings with regard to the prognostic gain using physical activity information in observational studies appears warranted before considering interventional studies comparing a physical activity-based risk stratification approach to risk stratification based on validated clinical models.

To summarise, in patients undergoing elevated-risk noncardiac surgery, self-reported functional capacity measures were independently associated with MACE. However, self-reported functional capacity information did not improve discrimination over a model based on clinical risk factors or over the NSQIP MICA. When using the RCRI as a risk estimation tool, self-reported functional capacity information statistically improved discrimination, but the effect size was limited and as such probably of limited clinical applicability. Limited self-reported functional capacity adversely influenced the risk of the major adverse cardiac events in patients submitted to noncardiac surgery. However, assessment of self-reported functional capacity expressed in METs or the other measured METs assessed here did not improve prediction of outcomes compared with clinical risk scores/clinical risk factors (i.e. it did not contribute to the differentiation of patients more likely to suffer an event). As such, caution should be applied in the use of self-reported functional capacity to guide clinical decisions in patients submitted to noncardiac surgery.

Strengths and limitations

The strengths of MET-REPAIR include a multicentre design across a large number of countries, a large number of events, and a population of patients at elevated cardiovascular risk (i.e. those patients in whom preoperative assessment of functional capacity is recommended).[2,3] Further, in addition to effort tolerance, we evaluated regular physical activity, as the last may provide a perspective different from maximal effort tolerance. This study has some limitations. First, whilst attending physicians were not actively informed of the answers to the questionnaire, there was a potential for performance bias because estimation of functional capacity for risk stratification is commonly included in preoperative assessment. However, the effect of preventive measures for cardiovascular events after noncardiac surgery is uncertain. Second,

the event rate for postoperative MI reported here is very low compared with previous studies. This is likely because of reduced sensitivity in endpoint detection because measurement of postoperative troponin was not mandated. A preliminary analysis of the various measures of functional capacity to predict myocardial injury suggested the same pattern (no improvement in prediction) as for the outcomes evaluated here. The results on myocardial injury are not presented here because (i) myocardial injury was not a primary endpoint; (ii) the centres applied various definitions in terms of age, comorbidities, and type of surgery to define the target population for the troponin surveillance; and (iii) the various assays used (conventional and high sensitivity). Third, although we encouraged the centres to include the most eligible patients and they captured 86% of eligible patients, we cannot exclude some non-responder bias. Of note, we did not ask centres to collect information on which exclusion criterion triggered exclusion, except for inability or unwillingness to participate. Fourth, the primary endpoint was a composite of major adverse cardiac events that had varying incidences and was mainly driven by cardiac death and MI. Further, in the perioperative literature, the components included in the composite MACE are inconsistent across studies, and expert consensus on what components MACE should include was published in October 2020[39] (i.e. after the MET-REPAIR study completed recruitment). Therefore, the MACE definition used here is not the one preferred by the Standardized Endpoints in Perioperative Medicine initiative.[39] Fifth, whilst the study detected a significant association between questionnaire-assessed METs dichotomised at 4 with MACE, it is important to specify that questionnaire-assessed 4 METs might not be equivalent to 4 METs measured during cardiopulmonary exercise testing but nearer to 5 METs (measured METs=0.51 × questionnaire-assessed METS+2.93).[9]

Further studies may address how other approaches to risk assessment (e.g. cardiac biomarkers) compare with self-reported functional capacity for the prediction of cardiovascular adverse events and in how far different approaches to preoperative risk assessment modify outcome.

Conclusions

In patients undergoing elevated-risk noncardiac surgery, self-reported functional capacity measures were independently associated with MACE. However, self-reported functional capacity information did not improve discrimination over a model based on clinical risk factors or over the NSQIP MICA. When using the RCRI as a risk estimation tool, self-reported functional capacity information statistically improved discrimination, but the effect size was limited and as such probably of limited clinical applicability. Limited self-reported functional capacity adversely influenced the risk of the major adverse cardiac events in patients submitted to noncardiac surgery. However, assessment of self-reported functional capacity expressed in METs or the other measured METs assessed here did not improve prediction of outcomes compared with clinical risk scores/clinical risk factors (i.e. it did not contribute to the differentiation of patients more likely to suffer an event). As such, caution should be applied in the use of self-reported functional capacity to guide clinical decisions resulting from risk estimation in patients submitted to noncardiac surgery.

Data availability statement

Requests for sharing of anonymised data for individual-level meta-analyses are to be addressed to the sponsor and steering committee and will be requested to follow the rules outlined in the study protocol.

Authors' contributions

Study conception/design: GLB, EM, DI, WS, SDH, MF, BBS, SJH
Data acquisition: GLB, EM, DI, WS, SDH, MF, BBS, SS, PM, DB, SCT, JvW, FL, KT, AG, H-JG, LG, KK, HW, JL, DC, SJH
Data analysis: GLB, FC
Data interpretation: GLB, EM, DI, WS, SDH, MF, BBS, SS, PM, DB, SCT, JvW, FL, KT, AG, H-JG, LG, KK, HW, JL, DC, FC, SJH
Drafting of paper: GLB, SJH
Revising of paper critically for important intellectual content: EM, DI, WS, SDH, MF, BBS, SS, PM, DB, SCT, JvW, FL, KT, AG, H-JG, LG, KK, HW, JL, DC, FC, SJH
Final approval of paper: all authors
All authors agree to be accountable for all aspects of the work, thereby ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Declaration of interest

GLB participated to an advisory board on perioperative myocardial injury hosted by Roche Diagnostics. The other authors have no conflicts of interest.

Funding

European Society of Anaesthesiology and Intensive Care Clinical Trial Network (GLB); Swiss Society of Anaesthesiology and Resuscitation (GLB); National Institute of Academic Anaesthesia (SJH; NIAA WKR020160060) .

Acknowledgements

The funding sources did not have any role in study design; data collection, analysis, and interpretation; writing of paper, or in the decision to submit the paper for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2023.02.030.

References

1. Rose J, Weiser TG, Hider P, Wilson L, Gruen RL, Bickler SW. Estimated need for surgery worldwide based on prevalence of diseases: a modelling strategy for the WHO Global Health Estimate. *Lancet Glob Health* 2015; **3**: S13–20
2. Halvorsen S, Mehilli J, Cassese S, et al. ESC guidelines on cardiovascular assessment and management of patients undergoing non-cardiac surgery. *Eur Heart J* 2022; **43**: 3826–924
3. Fleisher LA, Fleischmann KE, Auerbach AD, et al. ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac

surgery: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines. *J Am Coll Cardiol* 2014; **64**: e77–137
4. Wijeysundera DN, Pearse RM, Shulman MA, et al. Assessment of functional capacity before major non-cardiac surgery: an international, prospective cohort study. *Lancet* 2018; **391**: 2631–40
5. Lurati Buse GAL, Puelacher C, Gualandro DM, et al. Association between self-reported functional capacity and major adverse cardiac events in patients at elevated risk undergoing noncardiac surgery: a prospective diagnostic cohort study. *Br J Anaesth* 2021; **126**: 102–10
6. Marsman M, van Waes JAR, Grobben RB, Weersink CSA, van Klei WA. Added value of subjective assessed functional capacity before non-cardiac surgery in predicting postoperative myocardial injury. *Eur J Prev Cardiol* 2021; **28**: 262–9
7. Wijeysundera DN, Beattie WS, Hillis GS, et al. Integration of the Duke Activity Status Index into preoperative risk evaluation: a multicentre prospective cohort study. *Br J Anaesth* 2020; **124**: 261–70
8. Wiklund RA, Stein HD, Rosenbaum SH. Activities of daily living and cardiovascular complications following elective, noncardiac surgery. *Yale J Biol Med* 2001; **74**: 75–87
9. Jaeger C, Burkard T, Kamber F, et al. Quantification of metabolic equivalents (METs) by the MET-REPAIR questionnaire: a validation study in patients with a high cardiovascular burden. *J Clin Anesth* 2021; **76**, 110559
10. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation* 1999; **100**: 1043–9
11. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation* 2011; **124**: 381–7
12. Kristensen SD, Knuuti J, Saraste A, et al. ESC/ESA guidelines on non-cardiac surgery: cardiovascular assessment and management: the joint task force on non-cardiac surgery: cardiovascular assessment and management of the European Society of Cardiology (ESC) and the European Society of Anaesthesiology (ESA). *Eur J Anaesthesiol* 2014; **31**: 517–73
13. Thygesen K, Alpert JS, Jaffe AS, et al. Third universal definition of myocardial infarction. *J Am Coll Cardiol* 2012; **60**: 1581–98
14. Writing Committee for the VISION Study Investigators, Devereaux PJ, Biccard BM, et al. Association of postoperative high-sensitivity troponin levels with myocardial injury and 30-day mortality among patients undergoing noncardiac surgery. *JAMA* 2017; **317**: 1642–51
15. Ainsworth BE, Haskell WL, Herrmann SD, et al. Compendium of physical activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011; **43**: 1575–81
16. Holtermann A, Marott JL, Gyntelberg F, et al. Self-reported cardiorespiratory fitness: prediction and classification of risk of cardiovascular disease mortality and longevity—a prospective investigation in the Copenhagen City Heart Study. *J Am Heart Assoc* 2015; **4**, e001495
17. Stamatakis E, Hamer M, O'Donovan G, Batty GD, Kivimaki M. A non-exercise testing method for estimating cardiorespiratory fitness: associations with all-cause and cardiovascular mortality in a pooled analysis of eight population-based cohorts. *Eur Heart J* 2013; **34**: 750–8
18. Vascular Events In Noncardiac Surgery Patients Cohort Evaluation Study Investigators, Devereaux PJ, Chan MT, et al. Association between postoperative troponin levels and 30-day mortality among patients undergoing noncardiac surgery. *JAMA* 2012; **307**: 2295–304
19. Cohn SL, Fernandez Ros N. Comparison of 4 cardiac risk calculators in predicting postoperative cardiac complications after noncardiac operations. *Am J Cardiol* 2018; **121**: 125–30
20. Ford MK, Beattie WS, Wijeysundera DN. Systematic review: prediction of perioperative cardiac complications and mortality by the revised cardiac risk index. *Ann Intern Med* 2010; **152**: 26–35
21. Glance LG, Faden E, Dutton RP, et al. Impact of the choice of risk model for identifying low-risk patients using the 2014 American College of Cardiology/American Heart Association perioperative guidelines. *Anesthesiology* 2018; **129**: 889–900
22. Devereaux PJ, Mrkobrada M, Sessler DI, et al. Aspirin in patients undergoing noncardiac surgery. *N Engl J Med* 2014; **370**: 1494–503
23. Devereaux PJ, Xavier D, Pogue J, et al. Characteristics and short-term prognosis of perioperative myocardial infarction in patients undergoing noncardiac surgery: a cohort study. *Ann Intern Med* 2011; **154**: 523–8
24. Pearse RM, Moreno RP, Bauer P, et al. Mortality after surgery in Europe: a 7 day cohort study. *Lancet* 2012; **380**: 1059–65
25. Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *J Clin Epidemiol* 2001; **54**: 979–85
26. Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. *Biom J* 2005; **47**: 458–72
27. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007; **16**: 219–42
28. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med Res Methodol* 2017; **17**: 162
29. Kristensen SD, Knuuti J, Saraste A, et al. ESC/ESA guidelines on non-cardiac surgery: cardiovascular assessment and management: the joint task force on non-cardiac surgery: cardiovascular assessment and management of the European Society of Cardiology (ESC) and the European Society of Anaesthesiology (ESA). *Eur Heart J* 2014; **35**: 2383–431
30. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–45
31. Bouwmeester W, Moons KG, Kappen TH, et al. Internal validation of risk models in clustered data: a comparison of bootstrap schemes. *Am J Epidemiol* 2013; **177**: 1209–17
32. Wahl S, Boulesteix AL, Zierer A, Thorand B, van de Wiel MA. Assessment of predictive performance in incomplete data by combining internal validation and multiple imputation. *BMC Med Res Methodol* 2016; **16**: 144
33. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a

literature review and clinician's guide. *Ann Intern Med* 2014; **160**: 122–31
34. Pencina MJ, Steyerberg EW, D'Agostino Sr RB. Net reclassification index at event rate: properties and relationships. *Stat Med* 2017; **36**: 4455–67
35. Thomas LE, O'Brien EC, Piccini JP, D'Agostino RB, Pencina MJ. Application of net reclassification index to non-nested and point-based risk prediction models: a review. *Eur Heart J* 2019; **40**: 1880–7
36. World Health Organization. WHO guidelines on physical activity and sedentary behaviour 2020. Available from https://apps.who.int/iris/rest/bitstreams/1315866/retrieve. [Accessed 5 May 2021].
37. O'Donovan G, Stensel D, Hamer M, Stamatakis E. The association between leisure-time physical activity, low HDL-cholesterol and mortality in a pooled analysis of nine population-based cohorts. *Eur J Epidemiol* 2017; **32**: 559–66
38. Cortes OL, Moreno K, Alvarado P, Povea C, Lloyd M, Dennis R. Inactivity and its associated factors in adults scheduled for noncardiac surgery: the PAMP phase I study. *Rehabil Nurs* 2018; **43**: 81–7
39. Beattie WS, Lalu M, Bocock M, et al. Systematic review and consensus definitions for the Standardized Endpoints in Perioperative Medicine (StEP) initiative: cardiovascular outcomes. *Br J Anaesth* 2021; **126**: 56–66

Handling editor: Hugh C Hemmings Jr