Check for updates

Annals of Noninvasive Electrocardiology

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Examining the Effects of Different Exertion Levels and Reference Standards on the Diagnostic Accuracy of Exercise ECG for Detecting Coronary Artery Disease — A Pilot Study

Julius Heimann[1,2] | Thomas Schramm[3] | Martin Grauduszus[4] | Christine Joisten[2,4]

[1]Evangelisches Krankenhaus Duesseldorf, Dusseldorf, Germany | [2]University of Cologne, Faculty of Medicineand University Hospital Cologne, Cologne, Germany | [3]Kliniken der Stadt Koeln, Koln, Germany | [4]German Sport University Cologne, Koln, Germany

Correspondence: Julius Heimann (heimann.julius@outlook.de)

Received: 31 July 2024 | **Revised:** 10 November 2024 | **Accepted:** 19 November 2024

Funding: The authors received no specific funding for this work.

Keywords: coronary artery disease | diagnostic accuracy | exercise ECG | exertion | reference standard

ABSTRACT

Background: Exercise electrocardiogram (ECG) is frequently used for evaluating patients with suspected coronary artery disease (CAD). However, it has been downgraded in international guidelines due to a limited sensitivity and specificity, compared with other non-invasive tests.

Methods: We hypothesized, that a sufficient exertion going beyond 85% of the age predicted maximal heart rate (APMHR), by considering the achieved level of workload can improve sensitivity and specificity of the exercise ECG. Secondly, the performance of exercise ECG, when evaluated against a functional reference standard has barely been investigated, although an improved diagnostic accuracy has been reported for other non-invasive tests in this case. Therefore, in this pilot study, a total of 147 patients without known CAD who underwent exercise ECG followed by coronary angiography within 90 days were retrospectively evaluated.

Results: The combined end point of 85% of APMHR and 75% of the predicted workload showed a slight but not statistically significant benefit over APMHR alone (66.7% vs. 58.1%; $p = 0.545$). The sensitivity of ST-segment depression was 51.3% for detecting anatomically defined CAD and 62.5% for functionally relevant CAD ($p = 0.378$).

Conclusions: Although the results of this pilot study lack statistical significance, the numeric differences clearly justify further research on this topic.

1 | Introduction

Coronary artery disease (CAD) remains the most common cause of death worldwide (Institute for Health Metrics and Evaluation (IHME) 2023). Central to the diagnosis of CAD is the Risk-Factor-weighted Clinical Likelihood (RF-CL), which can be determined on the basis of symptoms, age, sex, and cardiovascular risk factors (Vrints et al. 2024; Diamond and Forrester 1979; Genders et al. 2011). If the RF-CL is very low (< 5%), current guidelines recommend no further diagnostic workup; if the RF-CL is high (> 85%), the presence of CAD is assumed to be probable without the need for confirmative diagnostic tests. For patients with a low or moderate RF-CL (5%–85%), non-invasive examinations are recommended. These examinations include exercise ECG, stress echocardiography, stress magnetic resonance imaging, single-photon emission computed tomography (SPECT), positron emission tomography, and coronary computed tomography (Vrints et al. 2024; Diamond and Forrester 1979; Genders et al. 2011; Knuuti et al. 2020; German Medical Association 2022; Montalescot et al. 2013).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Annals of Noninvasive Electrocardiology* published by Wiley Periodicals LLC.

A distinction can be made between functional and anatomical procedures for both non-invasive and invasive examinations (Singh et al. 2020). Functional procedures involve the detection of myocardial ischemia, for example, by ST-segment changes on the ECG, wall motion abnormalities or flow impairment (Balfour Jr., Gonzalez, and Kramer 2017). Anatomical procedures define CAD by the degree of visible stenosis (Tanabe et al. 2020).

In the outpatient setting, exercise ECG is one of the most frequently used non-invasive examinations in CAD diagnostics owing to its wide availability and simple feasibility (Montalescot et al. 2013). However, current guidelines are questioning the value of ergometry. European guidelines recommend the procedure only for patients with a RF-CL from > 5%–15% (Vrints et al. 2024). This recommendation is justified by the limited sensitivity of exercise ECG, which was reported in a large meta-analysis as having a mean of 68% (Detrano, Gianrossi, and Froelicher 1989).

Research shows that during exercise ECG, one third of coronary artery stenoses are discovered only when the heart rate exceeds 85% of its age-predicted maximum (Jain et al. 2011). However, to date there is a lack of studies evaluating new thresholds for a sufficient exertion. We assumed that a combined threshold of heart rate and workload may improve the diagnostic accuracy by excluding patients that did not reach a sufficient physical exertion during exercise ECG.

Secondly, studies on the sensitivity of exercise ECG are based exclusively on the comparison with an anatomical reference standard, predominantly invasive coronary angiography (ICA) (Banerjee et al. 2012). However, other functional tests (e.g., scintigraphy or echocardiography) show better performance when functionally relevant stenoses are used as the reference standard instead of stenoses defined anatomically via the level of diameter reduction (Knuuti et al. 2018). Since there is no study evaluating the exercise ECG against the two different reference standards, we further hypothesized that there is a higher correlation between ST-segment depressions and functionally relevant stenoses, than between ST-segment depressions and anatomically defined CAD. In other terms, we assumed that the sensitivity and specificity of exercise ECG are higher than previously described in the literature, when the aim is to detect functionally relevant stenoses.

2 | Methods

2.1 | Patient Selection

In this pilot study, patients were retrospectively selected from the patient registry of a German cardiology center. Based on the median study population of the meta-analysis of Knuuti et al. we aimed for the inclusion of 150 patients. Patients who received coronary angiography before February 2022 (time of data collection) were retrogradely assessed for inclusion criteria. These included an exercise ECG that was carried out no more than 3 months prior to ICA and the presence of a fully completed medical history form with information regarding pre-existing conditions, medications, current and past smoking history, family history of cardiovascular events and current cardiorespiratory complaints.

Patients with complete or incomplete left bundle branch block, those with ST-segment depression on resting ECG, those with known CAD and those with previous infarction or coronary intervention were excluded.

As visualized in Figure 1, to reach a final study population of 150 patients, a total number of 594 consecutive coronary angiographies reaching back to August 2016 had to be analyzed. After data extraction had been completed, patients taking digitalis ($n = 3$) were subsequently excluded, leaving a final study population of 147 patients.

2.2 | Patient Data and Medical History

The following patient data were collected from the medical history forms and the digital patient information system: age; sex; height; weight; current medication taken; nicotine abuse (current or past smoking history); family history of cardiovascular events; pre-existing conditions (e.g., hypercholesterolemia, diabetes mellitus, and arterial hypertension); presence of typical or atypical angina pectoris or dyspnea. Patient data were used to calculate the pre-test probability (precursor of the RF-CL used in the 2019 ESC Guideline), for the presence of CAD using the reference table recommended by the European Society of Cardiology at the time of data collection (Knuuti et al. 2020; Juarez-Orozco et al. 2019).

2.3 | Ergometry

Exercise testing was always performed on the same cycle ergometer (ergoselect 200 type K with HF, ergoline GmbH, Bitz, Germany) in the upright sitting position and under continuous 12-lead ECG monitoring. After a 2-min rest period and measurement of the resting blood pressure, exercise was typically started at 50 W and increased by 25 W every 2 min according to the WHO protocol (Löllgen and Leyk 2018). For patients with expectable low peak performance, the protocol was started at 25 W, for patients with expectable high performance, it was started at 75 or 100 W. The patients' standard medication was continued regularly on the day of the examination; in particular, beta-blocker administration was not paused. The same commonly accepted ending criteria were used in all patients (Fletcher et al. 1990).

All ECGs were measured manually using the custo diagnostic software (custo med GmbH, Ottbrunn, Germany) (version 5.6.8). Downsloping and horizontal ST-segment depression of at least 0.1 mV in at least two contiguous leads and at least three consecutive beats were considered significant. Downsloping and upsloping ST-segments were defined as those with a minimum slope of 1 mV/s, all others were considered horizontal. ST-segment depression was measured at either 60 ms (for HR > 130 bpm) or 80 ms (for HR < 130 bpm) after the J point, which was defined as the transition of the QRS complex into the ST segment (Fletcher et al. 2013; Khattar, Senior, and Lahiri 1998).

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

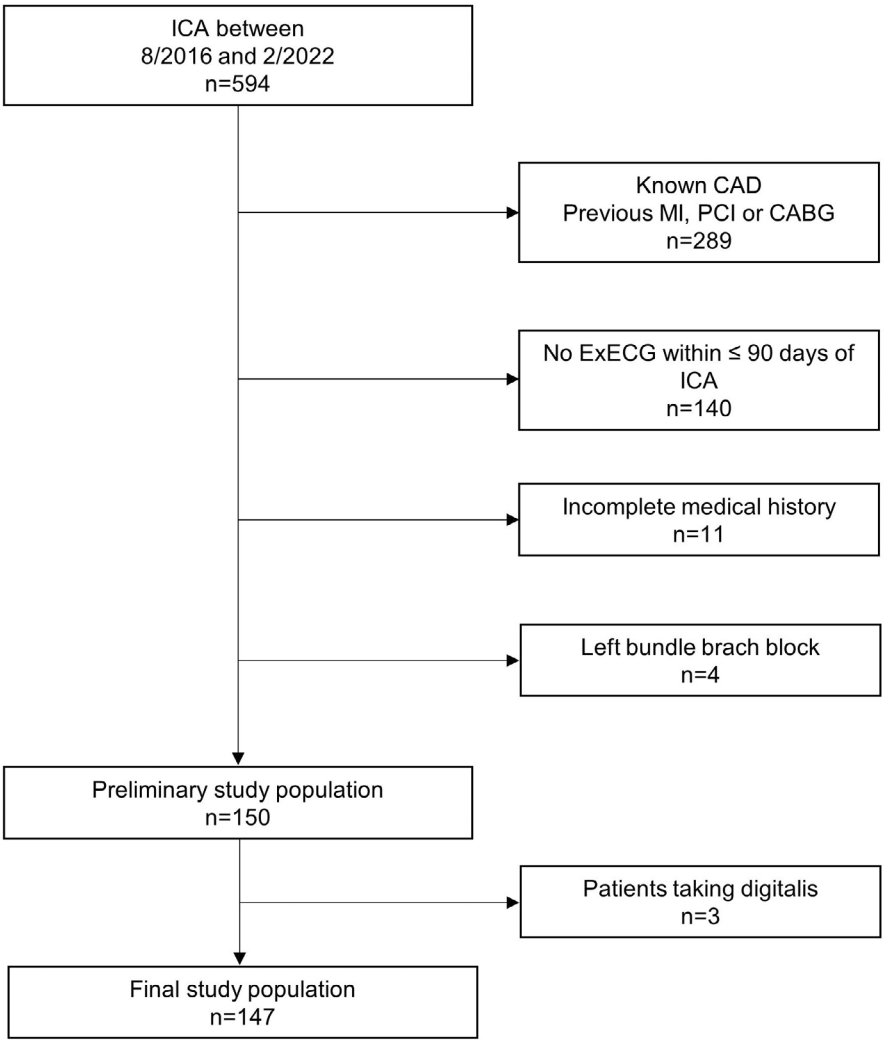

FIGURE 1 | Flowchart of patient selection. In total, 147 of 594 consecutive coronary angiographies were included in the final study population. The most common reasons for exclusion were known CAD or medical history of MI, PCI, or CABG.

The maximal heart rate and maximal achieved workload were set in relation to the age predicted maximal heart rate and the predicted workload for better comparability. The age predicted maximal heart rate was determined using the formula described by Tanaka, Monahan, and Seals (2001):

$$HR_{max} = 208 - (0.7 \times age)$$

The following formulas were used to determine the predicted workload (Wonisch et al. 2014): For males: Predicted workload (watts) = 6.773 + 136.141 × body surface area − 0.916 × body surface area × age. For females: Predicted workload (watts) = 3.933 + 86.641 × body surface area − 0.346 × body surface area × age.

The body surface area was calculated using the formula by Du Bois (1916):

$$\text{Body surface area } (m^2) = 0.007184 \times \text{height (cm)}^{0.725} \times \text{weight (kg)}^{0.425}$$

The patients were divided into a sufficient exertion group and an insufficient exertion group in three different scenarios. The criteria used were 85% of the maximal heart rate and/or 75% of the predicted workload. Accordingly, in scenario 1, the patients who reached either at least 85% of their APMHR or at least 75% of their predicted workload (or both) were considered as having exercised to a sufficient exertion level. In scenario 2, 85% of the APMHR alone were considered a sufficient exertion. In scenario 3, only patients who had reached both 85% of their APMHR and 75% of their predicted workload were considered to have a sufficient exertion.

2.4 | Coronary Angiography

All coronary angiographies were performed and evaluated by the same experienced interventional cardiologist. The investigator was not blinded to the outcome of the non-invasive diagnostic test. When at least one stenosis was present, the need for coronary intervention (stent or bypass operation) was assessed. Stenoses with less than 50% lumen reduction were considered functionally insignificant. Stenoses with more than 75% lumen reduction were considered to need revascularization. In the case of intermediate stenoses without evidence of ischemia from prior non-invasive testing FFR was used to assess hemodynamic relevance (Wijns et al. 2010).

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Patient characteristics of the intervention and no-intervention groups.

	All patients	No intervention	Intervention	
	$n = 147$	$n = 99$ (67.3%)	$n = 48$ (32.7%)	p
Age (year)	67.2 ± 10.0	67.2 ± 10.4	67.3 ± 9.2	0.95[a]
Female sex (%)	61 (41.5)	47 (47.5)	14 (29.2)	**0.04**[b]
Body mass index (kg/m^2)	28.6 ± 5.1	28.7 ± 4.6	28.4 ± 6.0	0.74[a]
PTP	22.9 ± 12.9	20.6 ± 11.6	28.4 ± 14.0	**< 0.01**[a]
PTP < 15%	45 (30.6)	36 (36.4)	9 (18.8)	**< 0.01**[b]
PTP > 85%	0	0	0	—
Current smoker	50 (34.0)	33 (33.3)	17 (35.4)	0.90[b]
Diabetes mellitus	28 (19.0)	15 (15.2)	13 (27.1)	0.08[b]
Hyperlipidemia	100 (68.0)	62 (62.6)	38 (79.2)	**0.04**[b]
Hypertension	112 (76.2)	74 (74.7)	38 (79.2)	0.56[b]
Positive family history of CVE	42 (28.6)	26 (26.3)	16 (33.3)	0.37[b]
Typical AP	40 (27.2)	22 (22.2)	18 (37.5)	0.05[b]
Intake of any of the beneath mentioned drugs:	129 (87.8)	85 (85.9)	44 (91.7)	0.31[b]
ß-blockers	62 (42.2)	44 (44.4)	18 (37.5)	0.42[b]
Calcium antagonists	24 (16.3)	15 (15.2)	9 (18.8)	0.58[b]
Potassium-saving diuretics	7 (4.8)	5 (5.1)	2 (4.2)	0.81[b]
Potassium diuretics	18 (12.2)	12 (12.1)	6 (12.5)	0.95[b]
Days from exercise ECG to ICA	26.4 ± 28.4	28.3 ± 29.0	22.6 ± 27.0	0.25[a]

Note: p-values in Bold: Significant difference ($p < 0.05$).
Abbreviations: AP, angina pectoris; CVE, cardiovascular event; ECG, electrocardiogram; ICA, invasive coronary angiography; PTP, pre-test probability.
[a]Unpaired *t*-test.
[b]Chi-square test.

2.5 | Statistical Analysis

All calculations were performed in IBM SPSS Statistics (IBM Corp, Armonk, NY, USA) (version 28.0.0.0 (190)) software. Four-fold contingency tables were used to determine sensitivity and specificity. Sensitivity was defined as the proportion of patients with ST-segment depressions out of all patients with a positive reference standard. Specificity was defined as the proportion of patients without ST-segment depressions out of all patients with a negative reference standard. For all continuous variables, mean values and standard deviations were calculated group-wise. An unpaired *t*-test was used to compare the mean values. Categorical variables are presented as absolute and relative frequencies and were compared using the Pearson chi-square test. Two-sided tails were selected as rejection areas and a p-value < 0.05 was considered significant.

3 | Results

3.1 | Demographic Data

Table 1 summarizes the characteristics of the participants. A total of 147 patients were included, of whom 61 were women and 86 were men. The mean patient age was 67.2 ± 10.0 years. Forty-eight patients (32.7%) required coronary intervention either by stenting or bypass surgery. This group requiring intervention ($n = 48$) did not differ from the group not needing intervention ($n = 99$) in age, BMI, smoking history, diabetes mellitus, hypertension rate, drug therapy, or positive family history of cardiovascular disease. In terms of pre-existing conditions, the only difference between the groups was the rate of dyslipidemia (79.2% in the intervention group vs. 62.6% in the no-intervention group, $p = 0.044$). There were significantly fewer women in the intervention group than in the no-intervention group (29.2% vs. 47.5%, $p = 0.035$); the pre-test probability was, on average, 7.8% greater in the intervention group than in the no-intervention group (28.4% vs. 20.6%, $p = 0.003$).

3.2 | Exercise Testing

Table 2 summarizes the maximal heart rate, maximal workload, blood pressure response, and events during exercise. The intervention and no-intervention groups did not differ significantly in the exercise duration, achieved heart rate or workload. Ventricular arrhythmias such as couplets, triplets, and ventricular tachycardia occurred rarely but equally in both groups. The resting systolic blood pressure was, on average, 6 mmHg higher in the intervention group than in the no-intervention group

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Exercise data, blood pressure response, and events during exercise.

	No intervention	Intervention	*p*
Exercise data			
Exercise duration (min)	6.9 ± 3.0	6.9 ± 3.7	0.99[a]
Maximal workload (W)	113.3 ± 45.9	119.1 ± 56.9	0.51[a]
Maximal workload in % of predicted workload	79.8 ± 25.8	78.1 ± 28.9	0.73[a]
HR_{max} reached (bpm)	130.6 ± 25.1	134.6 ± 26.0	0.37[a]
HR_{max} achieved in % of age predicted HR_{max}	81.2 ± 15.3	83.5 ± 15.2	0.39[a]
Achieved ≥ 85% of age predicted HR_{max}	39 (39.4%)	23 (47.9%)	0.33[b]
Achieved ≥ 75% of predicted workload	59 (59.6%)	26 (54.2%)	0.53[b]
Achieved 85% of HR_{max} AND 75% of predicted workload	31 (31.3%)	18 (37.5%)	0.46[b]
Achieved 85% of HR_{max} OR 75% of predicted workload	67 (67.7%)	31 (64.6%)	0.71[b]
Blood pressure response			
Resting RR_{sys} (mmHg)	141.5 ± 14.5	147.3 ± 17.4	**0.04**[a]
Resting RR_{dia} (mmHg)	84.8 ± 9.5	86.9 ± 9.3	0.22[a]
Maximal RR_{sys} during exercise (mmHg)	195.3 ± 29.7	202.7 ± 32.5	0.17[a]
Maximal RR_{dia} during exercise (mmHg)	95.6 ± 13.3	100.1 ± 14.8	0.09[a]
Events during exercise			
Angina pectoris	15 (15.2%)	12 (25.0%)	0.15[b]
Muscular fatigue limiting further exercise	83 (83.8%)	29 (60.4%)	**< 0.01**[b]
Dyspnea limiting further exercise	33 (33.3%)	11 (22.9%)	0.20[b]
Significant ST-segment depression	11 (11.1%)	22 (45.8%)	**< 0.001**[b]
≥ 1 Couplet	10 (10.1%)	4 (8.3%)	0.73[b]
≥ 1 Triplet	3 (3.0%)	0 (0.0%)	0.22[b]
Bigeminus	9 (9.3%)	3 (6.3%)	0.53[b]
Ventricular tachycardia	1 (1.0%)	1 (2.1%)	0.61[b]

Note: p-values in Bold: Significant difference ($p < 0.05$).
Abbreviations: HR_{max}, maximal heart rate; RR_{dia}, diastolic blood pressure; RR_{sys}, systolic blood pressure.
[a]Unpaired *t*-test.
[b]Chi-square test.

(147.3 vs. 141.5 mmHg, $p = 0.037$). While 83.8% of the patients in the no-intervention group discontinued the exercise test owing to muscular fatigue, this event was significantly less frequent (60.4%) among the patients with stenoses in the intervention group than in the no-intervention group ($p = 0.002$). Angina pectoris occurred in 25.0% of the patients in the intervention group and 15.2% in the no-intervention group, although this difference was not significant ($p = 0.148$).

The intervention group was significantly more likely to have horizontal or descending ST-segment depression of at least 0.1 mV on ECG than the no-intervention group (45.8% vs. 11.1%, $p < 0.001$). Since the sensitivity corresponds to the proportion of patients with a positive exercise ECG finding among all patients with a need for intervention, exercise ECG performed on a cycling ergometer had a sensitivity of 45.8% in our patient population based on all patients with a need for intervention and irrespective of sufficient exertion.

3.3 | Influence of Workload on Exercise ECG Sensitivity and Specificity

Table 3 summarizes the observed sensitivity and specificity of exercise ECG as a function of the workload criterion. In scenario 1, achieving one of the two parameters (85% of the maximal heart rate or 75% of the predicted workload) was considered sufficient exercise. This requirement was achieved by 31 of the 48 patients in the intervention group (64.6%). The most common criterion for sufficient exertion, which is 85% of the maximal heart rate without consideration of the workload (scenario 2), was achieved by 24 patients in the intervention group (50.0%). The strictest criteria (85% of the maximal heart rate and 75% of the predicted workload; scenario 3) were achieved by only 18 of the 48 patients in the intervention group (37.5%).

As depicted in Figure 2, the sensitivity was significantly higher among the patients with sufficient exertion in all three

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Sensitivity and specificity for ST-segment depression according to level of exertion.

	Achieved	Failed	Difference	p[a]	Effect size[b]
Scenario 1: 85% of HR_{max} and/or 75% of predicted workload					
All patients	76 (51.7%)	71 (48.3%)	—	—	—
Patients needing coronary intervention	31 (64.6%)	17 (35.4%)	—	—	—
With ST-segment depression (≙ sensitivity)	18 (58.1%)	4 (23.5%)	34.6%	**0.02**	0.33
Patients not needing coronary intervention	45 (45.5%)	54 (54.5%)	—	—	—
Without ST-segment depression (≙ specificity)	35 (77.8%)	53 (98.1%)	20.3%	**< 0.01**	0.29
Scenario 2: 85% of HR_{max}					
All patients	63 (49.2%)	84 (47.1%)	—	—	—
Patients needing coronary intervention	24 (50.0%)	24 (50.0%)	—	—	—
With ST-segment depression (≙ sensitivity)	15 (62.5%)	7 (29.2%)	33.3%	**0.02**	0.37
Patients not needing coronary intervention	39 (39.4%)	60 (60.6%)	—	—	—
Without ST-segment depression (≙ specificity)	29 (74.4%)	59 (98.3%)	23.9%	**< 0.001**	0.36
Scenario 3: 85% of HR_{max} and 75% of predicted workload					
All patients	32 (21.8%)	115 (78.2%)	—	—	—
Patients needing coronary intervention	18 (37.5%)	30 (72.5%)	—	—	—
With ST-segment depression (≙ sensitivity)	12 (66.7%)	10 (33.3%)	33.4%	**0.03**	0.32
Patients not needing coronary intervention	14 (14.1%)	85 (85.9%)	—	—	—
Without ST-segment depression (≙ specificity)	11 (78.6%)	77 (90.6%)	12.0%	0.19	0.26

Note: p-values in Bold: Significant difference ($p < 0.05$).
[a]Chi-square test.
[b]Cramer-V.

scenarios and at least twice as high as the sensitivity among the patients with insufficient exertion. Among the patients who did not meet the criteria for sufficient exertion in scenario 1 (heart rate or workload), the sensitivity was 23.5%; among the patients who met the criteria, the sensitivity was significantly higher at 58.1% ($p = 0.022$). Among the patients who did not reach the exertion criterion in scenario 2 (heart rate), the sensitivity was 29.2%; among the patients who reached the criterion, the sensitivity was significantly higher at 62.5% ($p = 0.020$). When the exertion criteria in scenario 3 (heart rate and workload) were not met, the corresponding patients showed a sensitivity of 33.3% compared with 66.7% among the patients who met the criteria ($p = 0.025$).

The sensitivity, as shown on the left side of Figure 3, increased from 58.1% when exertion criteria of scenario 1 were met, to 66.7% when exertion criteria of scenario 3 were met. However, this increase was not significant ($p = 0.545$). Among the respective non-exhausted patients, the sensitivity increased from 23.5% in scenario 1 to 33.3% in scenario 3, although this increase was also not significant ($p = 0.232$).

The specificity was consistently lower in the sufficient exertion group than in the insufficient exertion group, regardless of the exertion criterion used. For scenarios 1 and 2, this effect was significant ($p = 0.001$ and < 0.001, respectively). As shown on the right side of Figure 3, there was no visible trend for the change in specificity from scenarios 1 to 3.

The positive predictive value of a positive exercise ECG with sufficient exertion was 58.1%, 60.0% and 80.0% in scenarios 1, 2, and 3 respectively, without a statistical difference between the three scenarios (1 vs. 2: $p = 0.14$; 2 vs. 3: $p = 0.19$).

The negative predictive in scenarios 1–3 was 72.9%, 76.3%, and 64.7%, respectively (1 vs. 2: $p = 0.72$; 2 vs. 3: $p = 0.37$).

3.4 | Influence of the Reference Standard on Sensitivity and Specificity of Exercise ECG

Table 4 and Figure 4 describe the observed sensitivity and specificity according to the reference standard. Of the 147 patients, 48

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

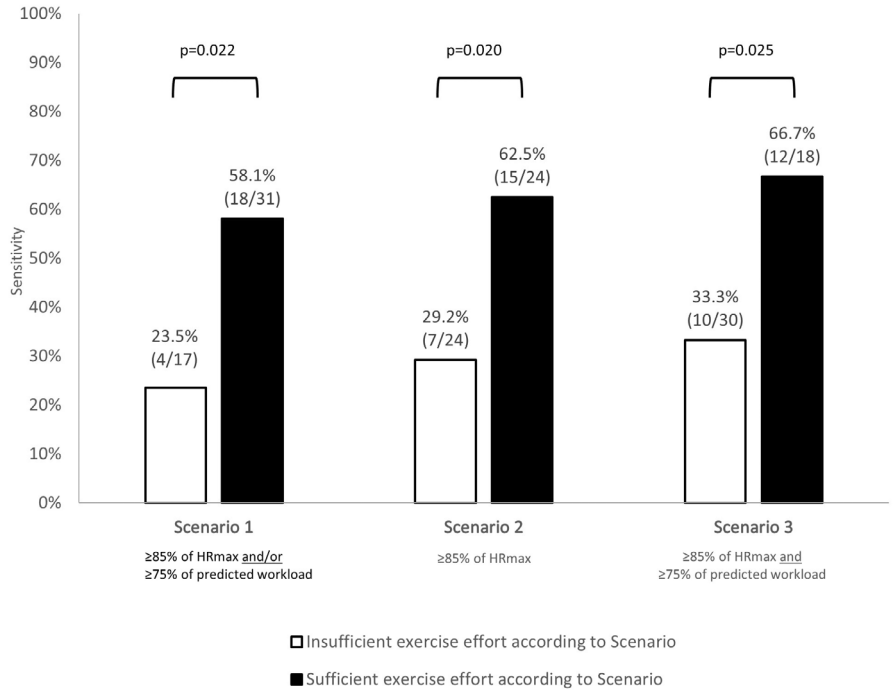

FIGURE 2 | Sensitivity of ST-segment depression on exercise ECG for the detection of coronary stenosis with need for coronary intervention ($n = 48$) for patients with sufficient and insufficient exertion, according to three different thresholds that define sufficient exertion.

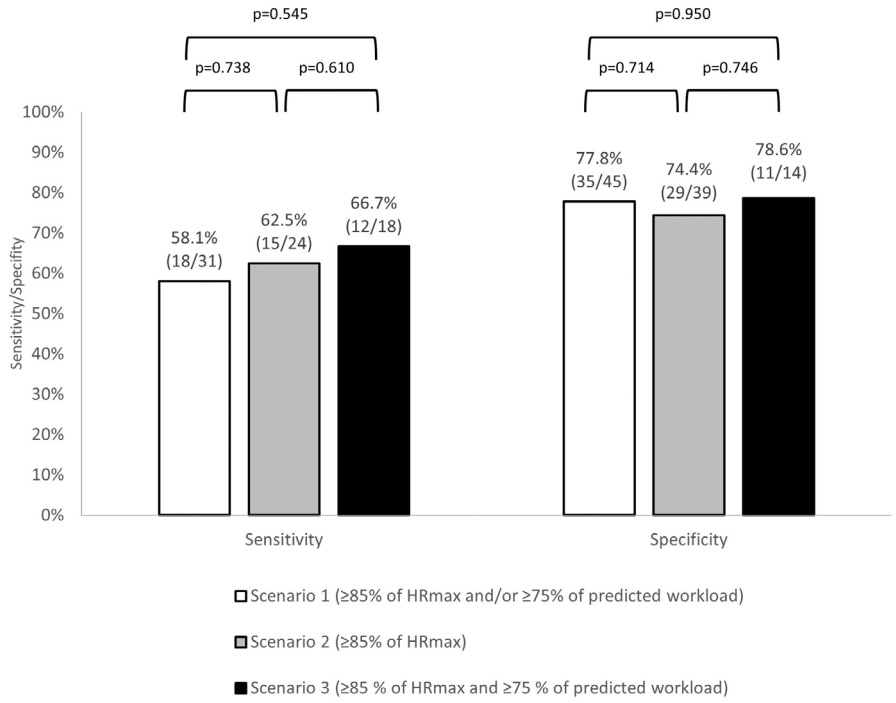

FIGURE 3 | Comparison of sensitivity and specificity for ST-segment depression in patients with sufficient exertion according to three different thresholds (Scenarios 1–3) that classify the exertion as sufficient or insufficient (Reference standard: Need for coronary intervention).

had one or more stenoses requiring intervention (32.7%). Eighty-three patients had ≥ 50% coronary stenosis (57.1%). We restricted the comparison of the reference standards to the patients who were exerting themselves based on their heart rate (scenario 2), as the heart rate represents the most common exertion criterion. Among 63 patients, 24 required intervention (38.1%) and 39 (61.9%) had ≥ 50% stenosis.

Of the 39 patients with sufficient exertion and a ≥ 50% stenosis, 20 developed significant ST-segment depression on exercise ECG. This corresponds to a sensitivity of 51.3%. Of the 24 patients with stenoses requiring intervention, 15 had significant ST-segment depression on exercise ECG, corresponding to a sensitivity of 62.5%. However, the difference of 11.2% in favor of the intervention-related reference standard was not significant

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Sensitivity and specificity for ST-segment depression according to the reference standard and achieved maximal heart rate.

	≥85% of HR_{max}	<85% of HR_{max}	Difference	*p*[a]	Effect size[b]
Patients needing coronary intervention	24	24	—	—	—
With ST-segment depression (≙ sensitivity)	15 (62.5%)	7 (29.2%)	33.3%	**0.02**	0.37
Patients not needing coronary intervention	39	60	—	—	—
Without ST-segment depression (≙ specificity)	29 (74.4%)	59 (98.3%)	23.9%	**<0.001**	0.36
Patients with at least 50% coronary stenosis	39	44	—	—	—
With ST-segment depression (≙ sensitivity)	20 (51.3%)	8 (18.2%)	33.1%	**<0.01**	0.38
Patients without at least 50% coronary stenosis	24	40	—	—	—
Without ST-segment depression (≙ specificity)	19 (79.2%)	40 (100%)	20.8%	**<0.01**	0.40

Note: p-values in Bold: Significant difference ($p < 0.05$).
[a]Chi-square test.
[b]Cramer-V.

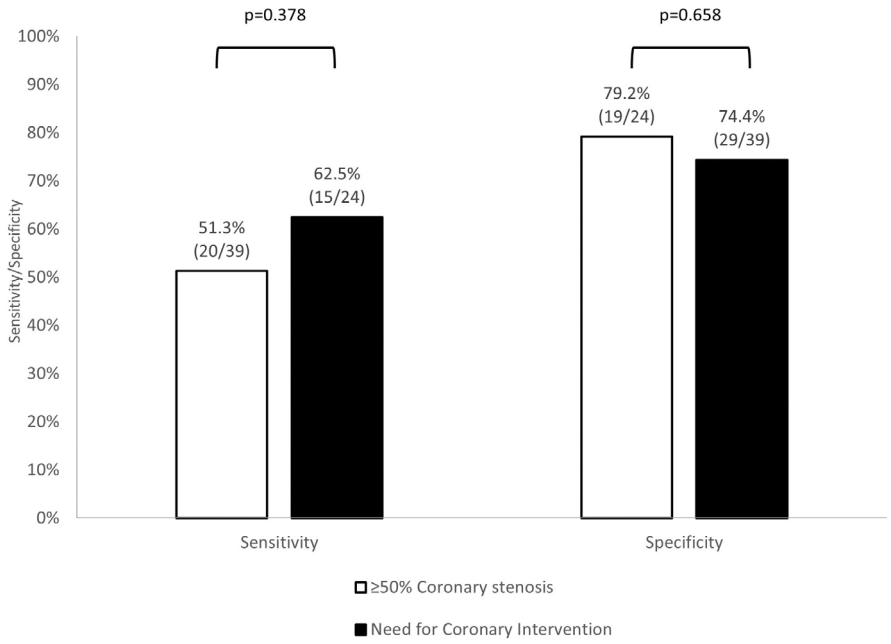

FIGURE 4 | Sensitivity and specificity of ST-segment depression according to the reference standard in patients that achieved at least 85% of the maximal age-predicted heart rate.

($p = 0.378$). The difference in sensitivity is visualized on the left side of Figure 4.

The specificity in relation to the reference standard was only slightly different and not significant (79.2% vs. 74.4%, $p = 0.658$). The difference in specificity is visualized on the right side of Figure 4.

The positive predictive value was 60.0% for the functional reference standard and 80.0% for the anatomic reference standard ($p = 0.12$). The negative predictive value was significantly higher for the functional reference standard (76.3%), compared with the anatomic reference standard (50.0%) ($p = 0.02$).

4 | Discussion

To date, few studies have evaluated the sensitivity and specificity of exercise ECG using different criteria to define sufficient exertion or a functional reference standard instead of

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the classical anatomic approach. Using a cohort of 147 patients who underwent exercise ECG at an adequate workload followed by coronary angiography within 90 days, we were able to document a sensitivity of 66.7% for the detection of stenoses requiring intervention. Achieving at least 85% of the maximal heart rate during exercise testing significantly improved the sensitivity from 29.2% to 62.5% in our study population ($p = 0.020$).

The importance of heart rate as an indicator of sufficient exertion is well established, which is the reason for the recommendation in the ergometry guidelines to use 85% of the maximal heart rate as a criterion for sufficient exertion (Fletcher et al. 2013; Klingenheben et al. 2018). Nevertheless, there were still nine cases of false-negative ECG findings among the patients who exercised according to this recommendation. In relation to the entire collective of 147 patients, the nine cases translate into 6.1% false-negative findings. This proportion could presumably be further reduced by adding the workload as an additional criterion for a sufficient exertion.

Although the threshold value of 85% is considered the minimum criterion for an interpretable stress test, it should not be misunderstood as being the termination criterion (Fletcher et al. 2013). In a study of over 300 patients, Jain et al. showed that one in three patients developed exercise-induced ST-segment depression only at more than 85% of their maximal heart rate (Jain et al. 2011). Especially in patients with chronotropic incompetence or those taking beta-blockers, the heart rate as the sole exertion criterion is of limited use.

Based on their clinical observations, Berent et al. recommended an additional minimum requirement of 75% of the predicted workload during ergometry (Berent et al. 2022). With the addition of this threshold value as a further exertion criterion (scenario 3), the number of exhausted patients in our cohort was reduced by 49.2% from 63 to 32 patients. Thus, in clinical practice, the proportion of patients requiring further diagnostic testing owing to inconclusive ergometry would increase. Of the 147 patients in our study, 115 (78.2%) would have had an inconclusive exercise test according to these criteria. Further, the rate of false negative findings, measured in all 147 patients examined, was reduced by the stricter exertion criteria, although not significantly, from 6.1% (scenario 2: heart rate only) to 4.1% (scenario 3: heart rate and workload; $p = 0.426$). The sensitivity increased, although also not significantly, from 62.5% to 66.7% ($p = 0.610$). Based on the above considerations of reducing false negatives and possibly increasing sensitivity, a stricter interpretation of the loading may be beneficial.

The second key point of our analysis was the comparison of the classical reference standard with a functional reference standard. Herein, we were able to document a sensitivity of 66.7% for the detection of stenoses requiring intervention. This contrasts with the sensitivity of 51.7% when using the classical reference standard ($\geq 50\%$ coronary stenosis). Consecutively the negative predictive value (NPV) of a normal exercise ECG rose from 50% (classical reference standard) to 76.3% (functional reference standard), leading to a significant improvement of the NPV by 26.3% ($p = 0.02$). Therefore, it can be speculated that sensitivity and the NPV of exercise ECG are underestimated when only coronary angiography is used as the reference standard.

A large meta-analysis by Knuuti et al. (2018) concluded that functional examinations (single photon emission computed tomography (SPECT), stress echocardiography and cardiac magnetic resonance (CMR)) have a significantly greater diagnostic power when the FFR, instead of conventional coronary angiography, is used as the reference standard. This is because whether a stenosis is functionally relevant is not determined by its diameter alone. Factors such as length, location and existing collateral circulation influence whether the stenosis leads to a relevant reduction in blood flow (Sun et al. 2015). Only stenoses that limit the flow to such an extent that they reduce the coronary reserve to a relevant degree can be detected on exercise ECG. If, however, the stenosis is considered only in relation to the reduction of the vessel diameter, it is questionable to what extent the visible lumen reduction correlates with the flow limitation.

In a study of over 600 patients, only 43% of stenoses that were visually assessed as high-grade showed a need for intervention, which was defined as an FFR of < 0.80 (Park et al. 2012). Knuuti et al. also described that coronary angiography itself has a sensitivity and specificity of only 68% and 73%, respectively, for detecting functionally significant stenoses (Knuuti et al. 2018).

If coronary angiography is therefore unable to reliably identify functionally relevant stenoses, the question inevitably arises as to what extent it can be a suitable reference standard for evaluating exercise ECG findings. For the question of whether coronary stenoses are present, regardless of their functional relevance, imaging methods such as coronary computed tomography are more suitable than exercise ECG. Exercise ECG rather helps determining the functional relevance of a given stenosis by providing evidence of myocardial ischemia. For this reason, we believe that exercise ECG should be evaluated against a functional reference standard.

There is also growing evidence that specificity of exercise ECG could be underestimated, because there is not only ischemia due to macrovascular obstructions, but also due to coronary microvascular dysfunction (Vrints et al. 2024). Sinha et al. examined patients with angina but nonobstructive coronary arteries. All patients that developed significant ST-depression during exercise ECG had evidence of coronary microvascular dysfunction. Therefore, they conclude that the high false positive rates that are described in the literature are due to a misunderstanding of myocardial ischemia (Sinha et al. 2024).

Our study has some limitations. For example, it is a retrospective pilot study in which only patients who had undergone coronary angiography were included. For ethical reasons, however, it would not be possible to subject healthy participants to ICA even in prospective studies. Furthermore, the investigator was not blinded to the outcome of the exercise test during coronary angiography. The result of the ergometric examination could therefore have influenced the assessment of the need for intervention. Beta-blockers were not paused on the day of the exercise ECG, which could have influenced the maximal heart rate. Finally, due to the limited sample size of 147 patients,

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

differences were statistically not significant, particularly in the subgroup analyses.

Larger studies should further investigate the different exertion criteria and reference standards to determine the added value of additional exertion criteria and to determine the true sensitivity of exercise ECG.

5 | Conclusion

Exercise ECG continues to play an important role in the assessment of patients with a cardiovascular risk profile in clinical practice, but is said to have limited sensitivity and specificity for the detection of CAD (Singh et al. 2020; Fletcher et al. 2013; Gibbons et al. 2002, German Medical Association 2019).

In our retrospective analysis of 147 patients performing bicycle ergometry tests followed by coronary angiography, we assessed the incremental impact of workload on the diagnostic accuracy of exercise ECG. The achievement of at least 85% of the maximal heart rate and 75% of the predicted workload in watts was associated with a non-significant 4.2% increase in sensitivity compared with the use of heart rate as the sole criterion for sufficient exertion. Therefore, it remains unclear whether a combination of heart rate and workload could lead to a superior diagnostic accuracy of exercise ECG.

Secondly, we compared the sensitivity and specificity of the exercise ECG when functionally relevant stenoses were used as the reference standard. Although not statistically significant, the sensitivity was 11.2% higher, when a functional, instead of an anatomical definition of CAD was used as the reference standard. It therefore seems reasonable to do further research on this topic.

Our pilot study has some important limitations. Beta-blockers were not withheld before exercise ECG. A selection bias applies, as only patient referred to a coronary angiography could be included. Finally, this pilot study possibly missed significant differences, due to its limited sample size.

Author Contributions

The ergometric and coronary angiographic examinations underlying this study were conducted by T.S. and his team. The concept of this study was developed by J.H. and C.S. equally. J.H. analyzed the ECGs and reviewed the coronary angiographic findings. Data collection, processing, as well as the statistical calculations were performed by J.H. Guidance on the selection of appropriate statistical methods was given by M.G. The manuscript was written by J.H. and subsequently reviewed by all co-authors, with C.J. making the most significant contributions to the review process.

Acknowledgments

The authors have nothing to report.

Ethics Statement

This is a retrospective study and does not require informed consent.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Balfour, P. C., Jr., J. A. Gonzalez, and C. M. Kramer. 2017. "Non-Invasive Assessment of Low- and Intermediate-Risk Patients With Chest Pain." *Trends in Cardiovascular Medicine* 27: 182–189. https://doi.org/10.1016/j.tcm.2016.08.006.

Banerjee, A., D. R. Newman, A. Van den Bruel, et al. 2012. "Diagnostic Accuracy of Exercise Stress Testing for Coronary Artery Disease: A Systematic Review and Meta-Analysis of Prospective Studies." *International Journal of Clinical Practice* 66: 477–492. https://doi.org/10.1111/j.1742-1241.2012.02900.x.

Berent, R., J. Auer, T. Berent, et al. 2022. *Ergometry 2022 - Role in Cardiovascular Medicine More Topical Than Ever (Ergometrie 2022 – Stellenwert in der kardiovaskulären Medizin Aktueller Denn Je).* JJATROS Kardiologie & Gefäßmedizin. https://www.universimed.com/ch/article/kardiologie-gefaessmedizin/ergometrie-stellenwert-medizin-142891.

Detrano, R., R. Gianrossi, and V. Froelicher. 1989. "The Diagnostic Accuracy of the Exercise Electrocardiogram: A Meta-Analysis of 22 Years of Research." *Progress in Cardiovascular Diseases* 32: 173–206. https://doi.org/10.1016/0033-0620(89)90025-X.

Diamond, G. A., and J. S. Forrester. 1979. "Analysis of Probability as an Aid in the Clinical Diagnosis of Coronary-Artery Disease." *New England Journal of Medicine* 300: 1350–1358. https://doi.org/10.1056/nejm197906143002402.

Du Bois, D. 1916. "A Formula to Estimate the Approximate Surface Area if Height and Weight Be Known." *Clinical Calorimetry* XVII: 863.

Fletcher, G. F., P. A. Ades, P. Kligfield, et al. 2013. "Exercise Standards for Testing and Training: A Scientific Statement From the American Heart Association." *Circulation* 128: 873–934. https://doi.org/10.1161/CIR.0b013e31829b5b44.

Fletcher, G. F., V. F. Froelicher, L. H. Hartley, W. L. Haskell, and M. L. Pollock. 1990. "Exercise Standards. A Statement for Health Professionals From the American Heart Association." *Circulation* 82: 2286–2322. https://doi.org/10.1161/01.CIR.82.6.2286.

Genders, T. S., E. W. Steyerberg, H. Alkadhi, et al. 2011. "A Clinical Prediction Rule for the Diagnosis of Coronary Artery Disease: Validation, Updating, and Extension." *European Heart Journal* 32: 1316–1330. https://doi.org/10.1093/eurheartj/ehr014.

German Medical Association (Bundesärztekammer) Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung), Association of the Scientific Medical Societies in Germany (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften). 2019. "National Guideline on Chronic Coronary Artery Disease (Nationale VersorgungsLeitlinie Chronische KHK) Version 5.0." https://doi.org/10.6101/AZQ/000419.

German Medical Association (Bundesärztekammer) Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung), Association of the Scientific Medical Societies in Germany (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften). 2022. "National Guideline on Chronic Coronary Artery Disease (Nationale VersorgungsLeitlinie Chronische KHK), Version 6.0." https://doi.org/10.6101/AZQ/000491.

Gibbons, R. J., G. J. Balady, J. T. Bricker, et al. 2002. "ACC/AHA 2002 Guideline Update for Exercise Testing: Summary Article. A Report of

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the American College of Cardiology/American Heart Association Task Force on Practice Guidelines Committee to Update the 1997 Exercise Testing Guidelines." *Journal of the American College of Cardiology* 40: 1531–1540. https://doi.org/10.1016/s0735-1097(02)02164-2.

Institute for Health Metrics and Evaluation (IHME). 2023. "GBD Compare." Accessed December 14, 2023. www.vizhub.healthdata.org/gbd-compare.

Jain, M., C. Nkonde, B. A. Lin, A. Walker, and F. J. T. Wackers. 2011. "85% of Maximal Age-Predicted Heart Rate Is Not a Valid Endpoint for Exercise Treadmill Testing." *Journal of Nuclear Cardiology* 18: 1026–1035. https://doi.org/10.1007/s12350-011-9454-0.

Juarez-Orozco, L. E., A. Saraste, D. Capodanno, et al. 2019. "Impact of a Decreasing Pre-Test Probability on the Performance of Diagnostic Tests for Coronary Artery Disease." *European Heart Journal Cardiovascular Imaging* 20: 1198–1207. https://doi.org/10.1093/ehjci/jez054.

Khattar, R. S., R. Senior, and A. Lahiri. 1998. "Assessment of Myocardial Perfusion and Contractile Function by Inotropic Stress Tc-99m Sestamibi SPECT Imaging and Echocardiography for Optimal Detection of Multivessel Coronary Artery Disease." *Heart* 79: 274–280. https://doi.org/10.1136/hrt.79.3.274.

Klingenheben, T., H. Löllgen, R. Bosch, and H. J. Trappe. 2018. "Manual on the Role of Ergometry (Manual Zum Stellenwert der Ergometrie)." *Cardiology (Kardiologie)* 12: 342–355. https://doi.org/10.1007/s12181-018-0265-2.

Knuuti, J., H. Ballo, L. E. Juarez-Orozco, et al. 2018. "The Performance of Non-invasive Tests to Rule-In and Rule-Out Significant Coronary Artery Stenosis in Patients With Stable Angina: A Meta-Analysis Focused on Post-Test Disease Probability." *European Heart Journal* 39: 3322–3330. https://doi.org/10.1093/eurheartj/ehy267.

Knuuti, J., W. Wijns, A. Saraste, et al. 2020. "2019 ESC Guidelines for the Diagnosis and Management of Chronic Coronary Syndromes." *European Heart Journal* 41: 407–477. https://doi.org/10.1093/eurheartj/ehz425.

Löllgen, H., and D. Leyk. 2018. "Exercise Testing in Sports Medicine." *Deutsches Ärzteblatt International* 115: 409–416. https://doi.org/10.3238/arztebl.2018.0409.

Montalescot, G., U. Sechtem, S. Achenbach, et al. 2013. "2013 ESC Guidelines on the Management of Stable Coronary Artery Disease: The Task Force on the Management of Stable Coronary Artery Disease of the European Society of Cardiology." *European Heart Journal* 34: 2949–3003. https://doi.org/10.1093/eurheartj/eht296.

Park, S.-J., S.-J. Kang, J.-M. Ahn, et al. 2012. "Visual-Functional Mismatch Between Coronary Angiography and Fractional Flow Reserve." *JACC: Cardiovascular Interventions* 5: 1029–1036. https://doi.org/10.1016/j.jcin.2012.07.007.

Singh, T., R. Bing, M. R. Dweck, et al. 2020. "Exercise Electrocardiography and Computed Tomography Coronary Angiography for Patients With Suspected Stable Angina Pectoris: A Post Hoc Analysis of the Randomized SCOT-HEART Trial." *JAMA Cardiology* 5: 920–928. https://doi.org/10.1001/jamacardio.2020.1567.

Sinha, A., U. Dutta, O. M. Demir, et al. 2024. "Rethinking False Positive Exercise Electrocardiographic Stress Tests by Assessing Coronary Microvascular Function." *Journal of the American College of Cardiology* 83: 291–299. https://doi.org/10.1016/j.jacc.2023.10.034.

Sun, L. J., G. S. Wang, M. Cui, et al. 2015. "Factors Influencing the Functional Significance in Intermediate Coronary Stenosis." *Journal of Geriatric Cardiology* 12: 107–112. https://doi.org/10.11909/j.issn.1671-5411.2015.02.008.

Tanabe, Y., A. Kurata, T. Matsuda, et al. 2020. "Computed Tomographic Evaluation of Myocardial Ischemia." *Japanese Journal of Radiology* 38: 411–433. https://doi.org/10.1007/s11604-020-00922-8.

Tanaka, H., K. D. Monahan, and D. R. Seals. 2001. "Age-Predicted Maximal Heart Rate Revisited." *Journal of the American College of Cardiology* 37: 153–156. https://doi.org/10.1016/S0735-1097(00)01054-8.

Vrints, C., F. Andreotti, K. C. Koskinas, et al. 2024. "ESC Guidelines for the Management of Chronic Coronary Syndromes: Developed by the Task Force for the Management of Chronic Coronary Syndromes of the European Society of Cardiology (ESC) Endorsed by the European Association for Cardio-Thoracic Surgery (EACTS)." *European Heart Journal* 45: 3415–3537. https://doi.org/10.1093/eurheartj/ehae177.

Wijns, W., P. Kolh, N. Danchin, et al. 2010. "Guidelines on Myocardial Revascularization: The Task Force on Myocardial Revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS)." *European Heart Journal* 31: 2501–2555. https://doi.org/10.1093/eurheartj/ehq277.

Wonisch, M., R. Berent, M. Klicpera, et al. 2014. "Practical Guidelines on Ergometry (Praxisleitlinien Ergometrie)." *Austrian Journal of Cardiology* 40: 40–244. https://doi.org/10.5414/ATX0947.

1542474x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/anec.70042, Wiley Online Library on [16/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License