

Comparative Evaluation of Treadmill-Based Risk Scores: Insights into Coronary Risk Stratification and Clinical Utility

✉ Hakan Gökulp Uzun, ✉ Burcu Uludağ Artun

University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Cardiology, İzmir, Türkiye

ABSTRACT

Introduction: Treadmill stress electrocardiogram (ECG) testing is widely used for coronary artery disease (CAD) assessment, but its accuracy is limited when based solely on ST-segment changes and exercise capacity. Several treadmill-based risk scores-Duke Treadmill score (DTS), Morise score (MS), Cleveland clinic prognostic score (CCPS), FIT Treadmill score (FITTS), and Rancho Bernardo score (RBS)-aim to improve risk stratification, but their comparative effectiveness remains unclear. This study evaluates their correlation and clinical applicability.

Methods: This cross-sectional study analyzed 136 patients undergoing treadmill stress testing at a cardiology outpatient clinic. Patients with contraindications or significant baseline ECG abnormalities were excluded. Demographic, clinical, and exercise test parameters were recorded. Risk scores were calculated using predefined equations, and Spearman's rank correlation was used to assess relationships among scores.

Results: The cohort had a mean age of 46 ± 13 years, with 43.4% women. Cardiovascular risk factors were common, including hyperlipidemia (24.3%), diabetes (11.8%), hypertension (25%), and smoking (43.4%). The categorical agreement was moderate between DTS and MS ($\kappa=0.42$) between CCPS and MS ($\kappa=0.37$), fair between MS and FITTS, DTS and RBS, and FITTS and CCPS ($\kappa=0.23-0.32$), and only slight for the remaining pairs ($\kappa=0.07-0.12$). Risk categorization varied significantly, with DTS and MS, predominantly classifying patients as low-risk, while FITTS and RBS provided a broader risk distribution.

Conclusion: Treadmill risk scores vary in the context of CAD risk classification. DTS is useful for identifying high-risk patients, while FITTS and CCPS may better assess lower-risk individuals. Combining scores may enhance risk stratification. Further research with long-term outcomes is needed.

Keywords: Treadmill stress testing, coronary artery disease, risk scoring models, Duke Treadmill score, Morise score, Cleveland clinic prognostic score, exercise capacity, risk stratification

Introduction

Coronary artery disease (CAD) is the leading cause of mortality and morbidity in the world (1). A timely diagnosis may prevent irreversible myocardial damage. Treadmill stress electrocardiogram (ECG) testing has been used as a primary tool for decades for the detection of CAD. The presence and/or degree of ST depression, along with the exercise intensity achieved during testing, provide some prognostic value; however, these variables have limited accuracy and precision. To enhance diagnostic and prognostic strength, various scoring models have been developed that incorporate additional variables beyond treadmill-based factors, including clinical and demographic risk factors.

Despite the proliferation of treadmill-derived algorithms, only a handful of studies have juxtaposed these scores directly and most were performed

more than two decades ago, enrolled highly selected male cohorts, or used heterogeneous endpoints such as angiographic stenosis versus clinical events (2-4). These observations underscore the importance of understanding how interchangeable traditional and contemporary treadmill scores, when applied to today's mixed-gender, risk-factor-rich outpatient population, are. Addressing this gap may help clinicians choose the most appropriate score for specific patient phenotypes and optimize downstream testing.

The optimal risk scoring scheme should be inclusive, incorporate all potential risk factors, and be easy to implement at the point of clinical care. Our study, based on a patient sample presenting to the cardiology outpatient clinic, aimed to compare the prominent treadmill score models and to investigate correlations as well as to illuminate their strengths and weaknesses.



Address for Correspondence: Hakan Gökulp Uzun MD, University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Cardiology, İzmir, Türkiye
E-mail: hakangokulpuzun@gmail.com ORCID ID: orcid.org/0000-0003-4194-0579

Cite this article as: Uzun HG, Uludağ Artun B. Comparative evaluation of treadmill-based risk scores: insights into coronary risk stratification and clinical utility. Istanbul Med J. 2025; 26(4): 259-64

Received: 17.03.2025

Accepted: 07.08.2025

Publication Date: 12.11.2025



©Copyright 2025 by the University of Health Sciences Türkiye, Istanbul Training and Research Hospital/Istanbul Medical Journal published by Galenos Publishing House.
Licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License

Methods

Study Population and Design

This is a descriptive, cross-sectional study designed to compare five different treadmill scores and explore how these scores vary based on the variables used in their calculation. The study cohort consisted of patients who underwent exercise ECG testing in our cardiology outpatient clinic with an indication for the diagnosis of coronary heart disease. Indications were selected at the discretion of the treating physician according to standard protocols in our outpatient clinic, following the latest guidelines on the subject. All consecutive applicants who were willing to comply with all study requirements were included. Accordingly, patients with any of the absolute contraindications for exercise stress test (EST) according to the guidelines were excluded (5). Furthermore, participants with baseline electrocardiographic abnormalities that might interfere with the assessment of ST-segment deviations, such as left bundle branch block or paced rhythm, were excluded. Patients underwent a comprehensive cardiac evaluation, including transthoracic echocardiography.

We collected demographic and clinical data including comorbidities and risk factors for coronary heart disease.

Test Procedure

EST was done using an integrated digital treadmill ECG system (GE T2100-ST, GE Healthcare, Chicago, Illinois, USA). Testing was performed following established guidelines (5). ECG recordings were obtained with a 12-lead system (Mason-Likar) with electrodes placed in modified positions (6).

Patients were instructed to continue their daily medications, including beta-blockers, as withholding does not appear to affect exercise performance (7).

The test was considered appropriate for the assessment if the patient reached 85% of their age-predicted maximum (max.) heart rate (APMHR) or achieved >7 metabolic equivalents (METs) of workload. APMHR was estimated using the equation “220-age” (8). The test was deemed “non-diagnostic” if neither of the two sufficiency criteria was fulfilled and the patient had no abnormal ECG changes.

Horizontal/downsloping ST-depression of at least 1 mm was considered abnormal (5). ST-segment depressions that were present before the exercise were subtracted from the peak depressions. The Treadmill scores (see Tables 1 and 2 for a summary).

Duke Treadmill Score (9)

The Duke Treadmill score (DTS) is the most widely used and cited exercise score since its invention in 1987. Its formula consists of three variables: total exercise time, largest ST-segment deviation in any lead measured in millimeters (except in lead aVR), and angina index (1= non-limiting angina and 2= exercise-limiting angina). It lacks some key variables such as age and heart rate. Total scores of $\geq +5$, -10 to +4, and < -10 correspond to low, intermediate, and high-risk levels, respectively, with associated 5-year survival rates of 99%, 95%, and 79%.

The equation is as follows: score = exercise time - [(5× ST-depression) + (4× angina index)].

Morise Score (Prognostic Exercise Test Scores for Men and Women) (10)

Developed in 2003 by Morise et al. (10), this externally validated tool differs from DTS mainly by separating scores by gender. The variables common to both genders are maximal heart rate, exercise ST-segment depression, age, angina history, diabetes, and the presence of exercise test-induced angina. Outside these shared domains, women are questioned about smoking and estrogen status, while men are asked

Table 1. Summary of Treadmill risk scores

Score	Year	Key variables	Risk categories	Survival estimates
Duke Treadmill score (9)	1987	- Total exercise time - ST-segment deviation (mm) - Angina index	• $\geq +5$: low-risk • -10 to +4: intermediate-risk • ≤ -10: high-risk	5-year survival: 99% (low) 95% (intermediate) 79% (high)
Morise score (10)	2003	- Maximal heart rate - Exercise ST-segment depression - Age - Angina history - Diabetes - Gender-specific (smoking, estrogen for women; hypercholesterolemia for men)	• < 40: low probability • 40-60: intermediate probability • > 60: high probability	Not specified
Cleveland clinic prognostic score (11)	2007	- Heart rate recovery - Frequent ventricular ectopy during recovery - Other clinical/test variables	Provides 3-, 5-, and 10-year survival estimates using an online tool	Varies by result
FIT treadmill score (12)	2015	- APPMHR - Maximum achieved workload (METs) - Age - Gender	Provides 10-year survival estimates based on a continuous total score	• > 100: 98% survival • 0 to 100: 97% survival • -1 to -100: 89% survival • ≤ -100: 62% survival
Rancho Bernardo (13)	2015	- ST-segment deviation (mm) - Not achieving target HR - Abnormal HR recovery - Chronotropic incompetency	Each abnormal response exponentially increases CHD and all-cause mortality risk	Not specified

APPMHR: Achieved percent of predicted maximum heart rate, CHD: Coronary heart disease, HR: Heart rate

about hypercholesterolemia. Each answer is assigned a point. Then all points are added together to get a total score. According to a total score, <40 points = low probability, 40-60 points = intermediate probability, and >60 points = high probability.

Cleveland Clinic Prognostic Score (11)

This scoring scheme includes variables not present in DTS or Morise score (MS), such as heart rate recovery and the presence of frequent ventricular ectopy during the recovery period. Available as an online tool (<https://riskcalc.org/SuspectedCoronaryArteryDiseaseLongTermSurvivalwNormalECG>), Cleveland clinic prognostic score (CCPS) provides estimates of 3-, 5- and 10-year survival based on clinical and test variables.

FIT Treadmill Score (12)

This score was derived from the 58,020 patients in the FIT project and provides estimates of all-cause mortality based on four simple variables: achieved percent of predicted max. heart rate (APPMHR), max. achieved workload as METs, patient age, and gender.

The equation is as follows: total score = APPMHR (%) +12x (METs) - 4x (Age) +43x (if female).

Scores greater than 100, 0 to 100, -1 to -100, and less than -100 gave 10-year median survival estimates of 98%, 97%, 89%, and 62%, respectively.

Rancho Bernardo Score (13)

According to the Rancho Bernardo score (RBS), each abnormal response to specific criteria increases the incidence of coronary heart disease and all-cause mortality exponentially. There is one electrocardiographic criterion: significant ST-change, defined as ST depression or elevation of 1 mm or more. Additionally, there are three non-electrocardiographic variables: not achieving the target heart rate, defined as at least 90% of the maximal heart rate predicted for age; abnormal heart rate recovery, defined as a drop of <22 bpm after 2 minutes of recovery; and

chronotropic incompetence, defined as failure to reach 80% of heart rate reserve.

Statistical Analysis

Variables are presented as mean $\pm$ standard deviation or median with interquartile range (IQR) for continuous variables and frequency (percent) for categorical variables. The normality of the variables was determined by the Kolmogorov-Smirnov test with a Lilliefors significance correction and the Shapiro-Wilk test. As the data had a nonparametric distribution, Spearman's rank correlation was used to assess how different risk scores correlate across patients. Each numeric risk score was stratified into low, intermediate, and high categories following their original publications. The agreement between categories was quantified with quadratic-weighted Cohen's κ , and a $\kappa \geq 0.61$ was considered substantial. Values were interpreted with the Landis and Koch scale.

The study was approved by the Non-Interventional Clinical Research Ethics Committee of University of Health Sciences Türkiye, İzmir City Hospital (approval number: 2024/186, date: 06.11.2024), and informed consent was obtained from all patients. Data analysis was conducted using the IBM SPSS Statistics software (version 26; IBM Corporation, Armonk, New York, United States).

Results

A total of 136 individuals were evaluated, with a mean age of 46 ± 13 years; 43.4% were women. On average, participants had a body mass index (BMI) of $26.8 (\pm 5.1)$, indicating an overweight profile. Notably, 24.3% had hyperlipidemia, 11.8% had diabetes, 47.8% had a family history of CAD, 25% had hypertension, and 43.4% were current or recent smokers (quit within the last year).

Regarding exercise test parameters, the mean resting heart rate was 86 ± 15.3 bpm, while the median peak heart rate was 159 bpm (IQR) 23, reflecting a significant chronotropic response. Participants' median functional capacity was 10 METs (IQR 2.6), indicating moderate-to-good

Table 2. Characteristics of the trials that developed the risk scores

Score	Authors	Year	Size (patients)	Characteristics	Validation	Follow-up	Age (years)	Notes
							Gender	
Duke Treadmill score	Mark et al. (15)	1987	2.842	Prospective All consecutive patients with anginal symptoms	External	Up to 10 yrs	49 (Mdn) 70% male	Most widely used yet lacks integration of patient features
Morise score	Morise et al. (10)	2003	4.640	Prospective All consecutive patients with anginal symptoms	External	2.8 ± 1.6 yrs	50 ± 12 53% male	One pretest score and two gender-specific exercise scores
Cleveland clinic prognostic score	Lauer et al. (11)	2007	33.268	Prospective Patients 30 years of age or older with anginal symptoms and a normal ECG	External	6.2 years (Mdn)	52 (Mdn) 62% male	Integrated analysis of ECG-related and non-ECG-related measures
FIT Treadmill score	Ahmed et al. (12)	2015	58.020	Retrospective All patients referred for an exercise stress test for any indication	Internal	10 years (Mdn)	18-96 (Mdn: 53) 51% male	Largest sample size
Rancho Bernardo	Shin et al. (13)	2015	1.789	Retrospective Asymptomatic adults	N/A	Up to 36 yrs	>20 (Mdn: 53) 47% male	Non-ECG variables had strong prognostic value

ECG: Electrocardiogram, Mdn: Median, N/A: Not applicable

exercise tolerance. The mean heart rate reserve-defined as the difference between resting and peak heart rate-was 62.9 ± 14.9 bpm, suggesting a generally preserved cardiovascular response among this middle-aged cohort. For a full look at the features of the cohort, refer to Table 3.

Most patients scored 0 on the DTS, indicating low-risk. Only a small proportion had a score of 1 or 3. A similar pattern was observed with CCPS, as almost all patients had a score of 0, and only a small number had a score of 3; indicating that most were at low-risk. There were a significant number of patients with a FIT Treadmill score (FITS) of 1 or 2, indicating a distribution of moderate to high-risk levels. Patients had predominantly low scores in MS, with very few showing high-risk. Finally, compared to the other scores, there was a wider distribution in RB, with more patients showing some risk. In brief, patients cluster in the low-risk band on DTS and CCPS, FITS and MS highlight more intermediate-risk patients, and RB is the most widely scattered (see Figure 1 for a breakdown of all scores).

MS and FITS showed the strongest association ($p=0.51$), suggesting these two scores share the most similarity in classifying patient risk. RBS was weakly correlated with both FITS ($p=0.05$) and CCPS ($p=0.09$). Finally, among the other four scores, DTS shows its strongest correlation with the RBS ($p=0.34$), suggesting a moderate relationship between these two risk stratification methods (refer to Figure 2 for a correlation matrix heatmap).

In the quadratic-weighted Cohen's κ agreement analysis, the strongest categorical concordance was between DTS and MS ($\kappa=0.42$), followed by CCPS-MS ($\kappa=0.37$) and FITS-CCPS ($\kappa = 0.32$). By contrast, the pair with the highest Spearman correlation coefficient in the earlier correlation analysis-MS-FITS ($p=0.51$)-showed only fair agreement ($\kappa=0.23$), underscoring that rank-order similarity does not guarantee consistent three-tier risk classification. DTS-RBS achieved fair agreement ($\kappa=0.25$) despite a moderate correlation ($p=0.34$), whereas RBS remained only slightly interchangeable with FITS ($p=0.05$; $\kappa=0.07$ -0.12) and with CCPS ($p=0.09$; $\kappa=0.07$ -0.12). all remaining score pairs exhibited slight agreement ($\kappa=0.07$ -0.12), highlighting the limited substitutability of these risk-stratification metrics (Table 4).

Table 3. Baseline characteristics

	Total
Variable	(n=136)
Demographics	
Age (years)	46±13
Women	59 (43.4%)
Clinical parameters	
Body mass index (kg/m²)	26.8 (5.1)
Hyperlipidemia	33 (24.3%)
Diabetes	16 (11.8%)
Family history of CAD	65 (47.8%)
Hypertension	34 (25%)
Smoking	59 (43.4%)
Exercise test parameters	
Resting heart rate (bpm)	86±15.3
Peak heart rate (bpm) [Mdn, (IQR)]	159 (23)
Functional capacity (METs) [Mdn, (IQR)]	10 (2.6)
Heart rate reserve	62.9±14.9

bpm: Beats per minute, CAD: Coronary heart disease, Mdn: Median, IQR: Interquartile range. Smoking was defined as being an active smoker, or having quit within the last year

Table 4. Agreement between risk scores

Pair (3-tier categories)	Quadratic-weighted κ	Landis-koch strength
DTS vs. MS	0.42	Moderate
CCPS vs. MS	0.37	Moderate
MS vs. FITS	0.23	Fair
DTS vs. RBS	0.25	Fair
FITS vs. CCPS	0.32	Fair
All others	0.07-0.12	Slight

CCPS: Cleveland clinic prognostic score, DTS: Duke Treadmill score, FITS: FIT Treadmill score, MS: Morise score, RBS: Rancho Bernardo score

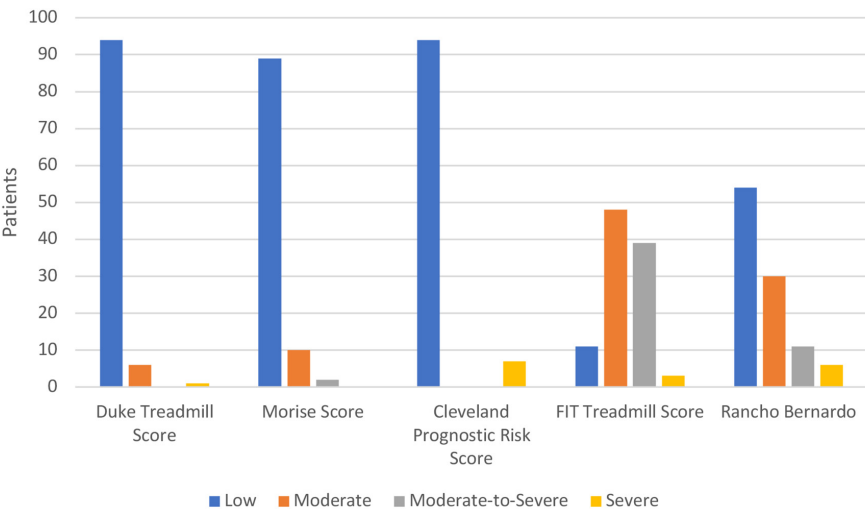
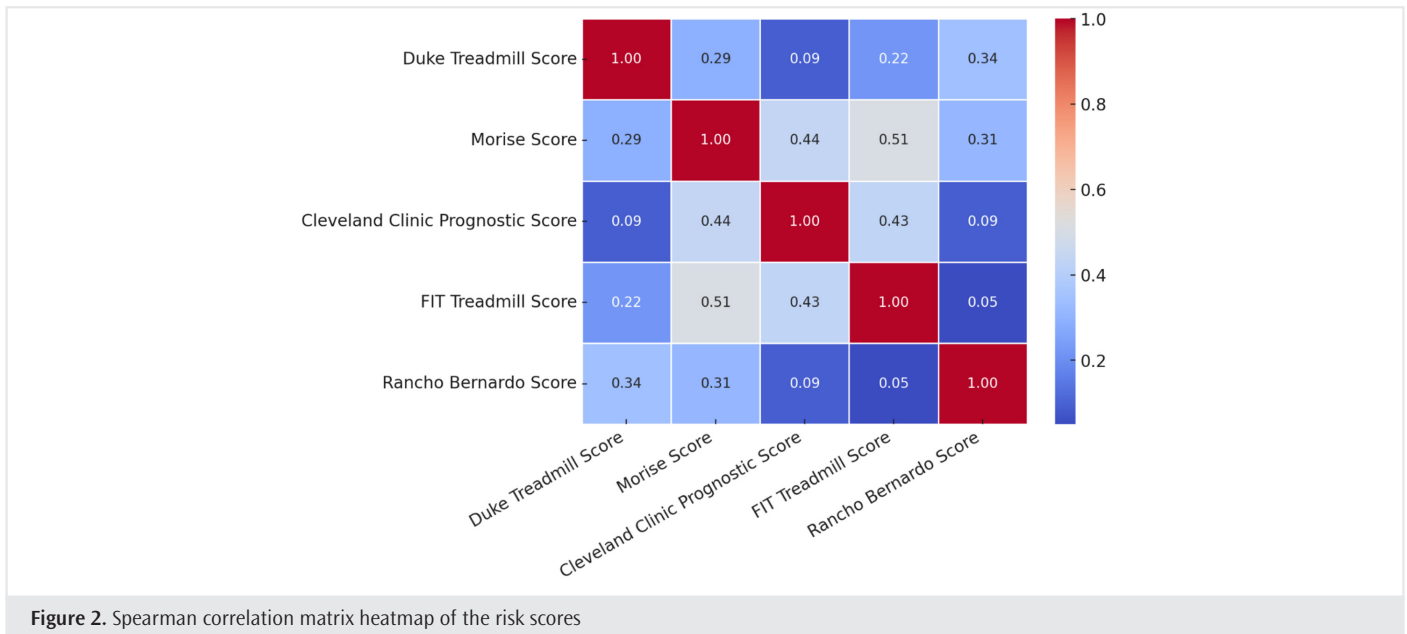


Figure 1. Distribution of the patients across five scores



Discussion

Despite strong rank correlations, agreement analysis showed only moderate concordance between certain scores and marginal agreement between others, underscoring that the scores are not interchangeable in individual patients. The loss of significance of association between MS and FITTS after further statistical analysis illustrates the fact that correlation measures parallel trends, whereas κ quantifies exact category matching. Thus, MS and FITTS, though paralleling each other ($p=0.51$), agree on the exact category only ~23 % better than chance.

Our finding that the MS and DTS exhibit the strongest categorical agreement is consistent with the large angiography-validated cohort analyzed by Fearon et al. (2), ($n=1,282$), in which both algorithms demonstrated comparable performance in detecting $\geq 50\%$ stenosis (0.77 ± 0.01 vs. 0.73 ± 0.01 , respectively). The DTS was developed to pinpoint high-risk individuals (those with a high pretest likelihood of coronary heart disease) by forecasting significant stenosis on invasive coronary angiography ($\geq 75\%$), thus aiding in determining when invasive angiography is warranted for patients with chest pain. However, in people who are lower-risk and have normal test findings-especially those without symptoms-the DTS provides limited additional benefit compared to simply assessing exercise capacity. Our cohort matched the low-risk profile described in the original Duke papers, with most individuals scoring ≥ 5 and only a few scoring 1 or 3. By contrast, the FITTS was created for lower-risk patients whose likelihood of coronary heart disease after testing remains low; this clarifies why the DTS and FITTS do not align closely.

The FITTS and the RBS tend to assign higher-risk values overall, each with a median value of 1, while the other three scores have median values of 0. This suggests, these two scores may be more sensitive to detecting potential risk factors. However, the RBS shows weaker correlations with most other scores, indicating it may assess distinct dimensions of risk. Although the EST positivity rate in the RB study was low (approximately

6%; $n=8$) - a rate consistent with other studies-non-ECG measures provided robust insights into the future risk of cardiovascular mortality.

Although the FITTS and MS scores demonstrated a moderate correlation ($p=0.51$) and fair agreement ($\kappa=0.23$), their concordance in categorizing patients was not proportionally high. This further underscores that correlation or agreement alone does not ensure consistent risk stratification, highlighting the necessity of evaluating categorical concordance when comparing risk scores.

A moderate correlation was observed between the CCPS and the MS ($\kappa=0.37$), suggesting a fair agreement in the risk categorization of patients. Both scores include extensive clinical variables, including heart rate improvement, exercise-induced angina, and other demographic and cardiovascular factors that likely contribute to their correlation. In the derivation cohort of the original CCPS study, 64% of patients identified as moderate or high-risk by DTS were reclassified as low-risk by CCPS. Nearly all patients in our cohort scored 0; this replicated the low-risk distribution also seen in its validation cohorts. This may explain the very low correlation between CCPS and DTS in our study. In summary, both CCPS and FITTS seem to have a good discriminative capacity for low- and intermediate-risk patients, while DTS would be an ideal choice for high-risk patients who are more likely to need more advanced invasive investigations such as coronary angiography.

Almost half (43.4%) of patients smoke, a quarter (25%) have hypertension, and BMI data show a trend toward overweight and obesity (Mdn 26.8). This high prevalence of modifiable risk factors means that the population would benefit from aggressive risk factor modification and preventive interventions.

These findings underscore the complementary nature of the risk scores. CCPS and MS are well suited to comprehensive assessment in data-rich clinical settings, whereas FITTS, which relies mainly on demographic factors and exercise capacity, is ideal for broad population screening. RBS adds targeted insight in special circumstances by emphasizing

non-ECG markers such as chronotropic incompetence and heart-rate recovery. Consequently, combining scores a priori may yield additive prognostic value, as each captures distinct pathophysiologic domains (e.g., chronotropic response versus METs). Clinicians should therefore select scores judiciously: CCPS can sharpen stratification in older adults or those with autonomic dysfunction, while FITS is preferable for gauging fitness-related risk in younger, otherwise healthy individuals.

Study Limitations

This study is limited by the lack of long-term clinical outcomes (e.g., mortality or cardiac adverse events) to confirm the predictive power of these scores. Additionally, we did not include a gold standard imaging comparator, such as invasive coronary angiography or coronary CT angiography, which limits our ability to comment on the absolute diagnostic accuracy of each score in our cohort. The observed discrepancies in correlation highlight the need for studies examining whether combining multiple risk scores-potentially augmented with artificial intelligence-can improve predictive accuracy. Emerging evidence likewise indicates that machine-learning-enhanced treadmill analytics may more precisely refine risk stratification (14). Further research is needed to investigate how these variables might complement existing scores or serve as the basis of new risk models.

Conclusion

This study highlights the variability and complementary nature of treadmill-based risk scores for CAD. While DTS is effective for identifying high-risk patients needing invasive evaluation, CCPS and FITS are better suited for lower-risk populations, and RBS adds value with non-ECG parameters. Moderate correlations and notable discrepancies among scores suggest that tailored selection enhances risk stratification. While the five treadmill-derived scores move broadly in the same direction, their modest κ values show they should not be used interchangeably for individual patient decisions. Future prospective studies should evaluate whether integrating complementary variables from more than one treadmill score can further refine clinical decision-making.

Ethics

Ethics Committee Approval: The study was approved by the Non-Interventional Clinical Research Ethics Committee of University of Health Sciences Türkiye, İzmir City Hospital (approval number: 2024/186, date: 06.11.2024).

Informed Consent: Informed consent was obtained from all patients.

Footnotes

Authorship Contributions: Surgical and Medical Practices - H.G.U., B.U.A.; Concept - H.G.U., B.U.A.; Design - H.G.U., B.U.A.; Data Collection or Processing - H.G.U., B.U.A.; Analysis or Interpretation - H.G.U., B.U.A.; Literature Search - H.G.U.; Writing - H.G.U.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. WHO. Global health estimates: leading causes of death. World Heal Organ. 2019.
2. Fearon WF, Gauri AJ, Myers J, Raxwal VK, Atwood JE, Froelicher VF. A comparison of treadmill scores to diagnose coronary artery disease. Clin Cardiol. 2002; 25: 117-22.
3. Alexander KP, Shaw LJ, Shaw LK, DeLong ER, Mark DB, Peterson ED. Value of exercise treadmill testing in women. J Am Coll Cardiol. 1998; 32: 1657-64.
4. Cuenza LR, Yap EML, Ebba E. Assessment of the prognostic utility of the FIT treadmill score in coronary artery disease patients undergoing cardiac rehabilitation. J Cardiovasc Thorac Res. 2019; 11: 8-13.
5. Fletcher GF, Ades PA, Kligfield P, Arena R, Balady GJ, Bittner VA, et al. Exercise standards for testing and training: a scientific statement from the American heart association. Circulation. 2013; 128: 873-934.
6. Mason RE, Likar I. A new system of multiple-lead exercise electrocardiography. Am Heart J. 1966; 71: 196-205.
7. Herbert WG, Dubach P, Lehmann KG, Froelicher VF. Effect of beta-blockade on the interpretation of the exercise ECG: ST level versus delta ST/HR index. Am Heart J. 1991; 122: 993-1000.
8. Fox SM 3rd, Naughton JP, Haskell WL. Physical activity and the prevention of coronary heart disease. Ann Clin Res. 1971; 3: 404-32.
9. Mark DB, Hlatky MA, Harrell FE Jr, Lee KL, Califf RM, Pryor DB. Exercise treadmill score for predicting prognosis in coronary artery disease. Ann Intern Med. 1987; 106: 793-800.
10. Morise AP, Jalisi F. Evaluation of pretest and exercise test scores to assess all-cause mortality in unselected patients presenting for exercise testing with symptoms of suspected coronary artery disease. J Am Coll Cardiol. 2003; 42: 842-50.
11. Lauer MS, Pothier CE, Magid DJ, Smith SS, Kattan MW. An externally validated model for predicting long-term survival after exercise treadmill testing in patients with suspected coronary artery disease and a normal electrocardiogram. Ann Intern Med. 2007; 147: 821-8.
12. Ahmed HM, Al-Mallah MH, McEvoy JW, Nasir K, Blumenthal RS, Jones SR, et al. Maximal exercise testing variables and 10-year survival: fitness risk score derivation from the FIT project. Mayo Clin Proc. 2015; 90: 346-55.
13. Shin SY, Park JI, Park SK, Barrett-Connor E. Utility of graded exercise tolerance tests for prediction of cardiovascular mortality in old age: the Rancho Bernardo study. Int J Cardiol. 2015; 181: 323-7.
14. Wei J, Pan B, Gan Y, Li X, Liu D, Sang B, et al. Temporal relationship-aware treadmill exercise test analysis network for coronary artery disease diagnosis. Sensors (Basel). 2024; 24: 2705.
15. Mark DB, Shaw L, Harrell FE Jr, Hlatky MA, Lee KL, Bengtson JR, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. N Engl J Med. 1991; 325: 849-53.