

Real-world evaluation of an algorithmic machine-learning-guided testing approach in stable chest pain: a multinational, multicohort study

Evangelos K. Oikonomou ¹, Arya Aminorroaya¹, Lovedeep S. Dhingra¹, Caitlin Partridge², Eric J. Velazquez ¹, Nihar R. Desai¹, Harlan M. Krumholz^{1,3}, Edward J. Miller ¹, and Rohan Khera ^{1,3,4,5,*}

¹Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, 333 Cedar Street, PO Box 208017, New Haven, 06520-8017 CT, USA; ²Yale Center for Clinical Investigation, 2 Church Street South, New Haven, 06519 CT, USA; ³Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, 195 Church Street 5th Floor, New Haven, 06510 CT, USA; ⁴Section of Biomedical Informatics and Data Science, Yale School of Medicine, 100 College Street, New Haven, 06511 CT, USA; and ⁵Section of Health Informatics, Department of Biostatistics, Yale School of Public Health, 60 College Street, New Haven, 06510 CT, USA

Received 3 January 2024; revised 13 February 2024; accepted 20 March 2024; online publish-ahead-of-print 8 April 2024

Aims

An algorithmic strategy for anatomical vs. functional testing in suspected coronary artery disease (CAD) (Anatomical vs. Stress teSting decision Support Tool; ASSIST) is associated with better outcomes than random selection. However, in the real world, this decision is rarely random. We explored the agreement between a provider-driven vs. simulated algorithmic approach to cardiac testing and its association with outcomes across multinational cohorts.

Methods and results

In two cohorts of functional vs. anatomical testing in a US hospital health system [Yale; 2013–2023; $n = 130\,196$ (97.0%) vs. $n = 4020$ (3.0%), respectively], and the UK Biobank [$n = 3320$ (85.1%) vs. $n = 581$ (14.9%), respectively], we examined outcomes stratified by agreement between the real-world and ASSIST-recommended strategies. Younger age, female sex, Black race, and diabetes history were independently associated with lower odds of ASSIST-aligned testing. Over a median of 4.9 (interquartile range [IQR]: 2.4–7.1) and 5.4 (IQR: 2.6–8.8) years, referral to the ASSIST-recommended strategy was associated with a lower risk of acute myocardial infarction or death (hazard ratio_{adjusted}: 0.81, 95% confidence interval [CI] 0.77–0.85, $P < 0.001$ and 0.74 [95% CI 0.60–0.90], $P = 0.003$, respectively), an effect that remained significant across years, test types, and risk profiles. In *post hoc* analyses of anatomical-first testing in the Prospective Multicentre Imaging Study for Evaluation of Chest Pain (PROMISE) trial, alignment with ASSIST was independently associated with a 17% and 30% higher risk of detecting CAD in any vessel or the left main artery/proximal left anterior descending coronary artery, respectively.

Conclusion

In cohorts where historical practices largely favour functional testing, alignment with an algorithmic approach to cardiac testing defined by ASSIST was associated with a lower risk of adverse outcomes. This highlights the potential utility of a data-driven approach in the diagnostic management of CAD.

* Corresponding author. Tel: +1 203 764 5885, Email: rohan.khera@yale.edu, @rohan_khera

© The Author(s) 2024. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Structured Graphical Abstract

Key Question

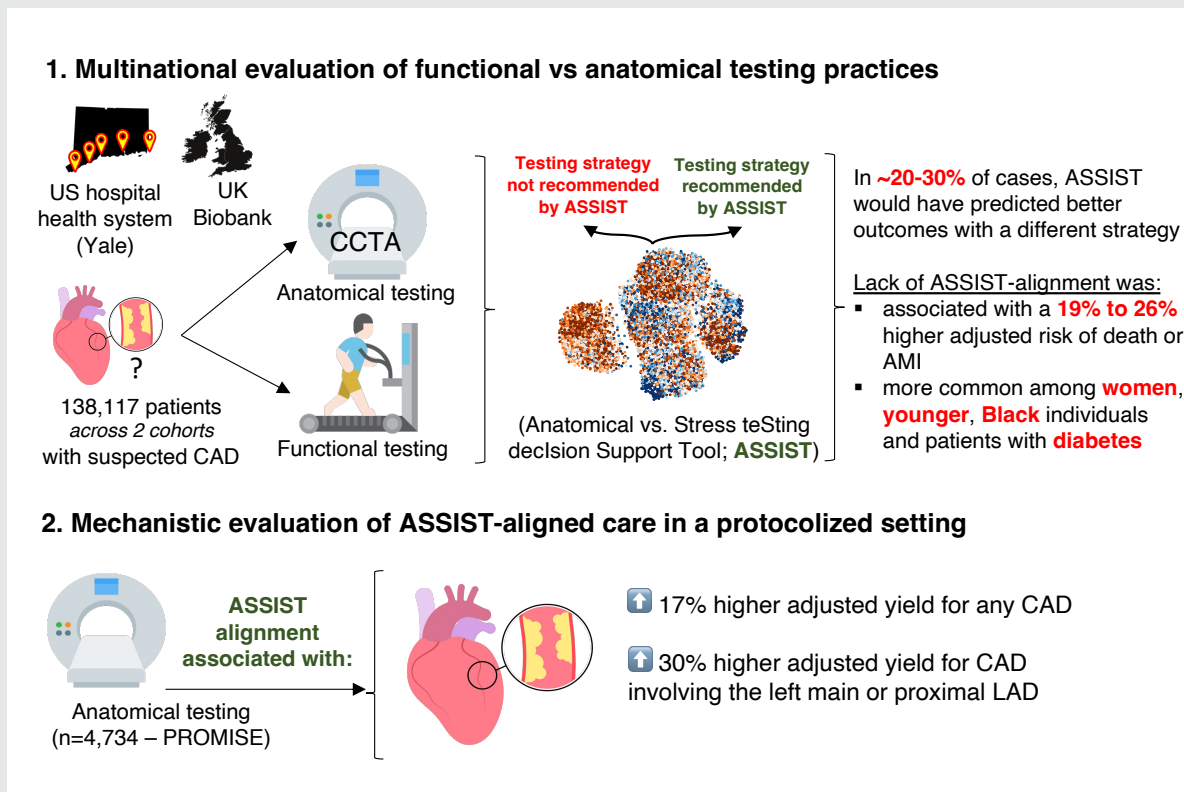
Data-driven test selection using a machine-learning-derived algorithmic strategy (Anatomical vs. Stress teSting decision Support Tool, ASSIST) for anatomical vs. functional testing in stable chest pain was associated with better outcomes than random selection of testing strategy based on data drawn from clinical trials. However, it is unclear whether this association is present in the real world where test selection is already driven by clinician- and patient-specific factors.

Key Finding

Across two multinational cohorts, including a large US-based hospital system and the UK Biobank, ASSIST-aligned testing was reproducibly associated with a 19–26% lower risk of death or acute myocardial infarction. Younger individuals, female patients, and racial minorities were less likely to undergo ASSIST-recommended testing. In a mechanistic substudy in PROMISE, alignment with an ASSIST recommendation for anatomical testing was associated with higher anatomical yield.

Take Home Message

Real-world alignment with an algorithmic approach to functional vs. anatomical testing is independently associated with better long-term outcomes, suggesting the potential utility of an algorithmic approach to cardiac testing.



Keywords Machine learning • Chest pain • Artificial intelligence • Clinical decision support

Introduction

The workup of stable chest pain and suspected coronary artery disease (CAD) relies on a range of non-invasive tools to refine the initial risk stratification achieved through history and physical examination. In the modern era, these generally include functional stress tests and anatomical imaging by coronary computed tomography angiography (CCTA).^{1–3} The choice between the two approaches often depends on patient-specific factors, availability, and local expertise,^{1,3–5} with commonly cited sensitivity and specificity metrics calibrated against the cross-sectional presence of obstructive CAD on invasive angiography,⁶ rather than long-term cardiovascular outcomes. A head-to-head comparison of

a functional- vs. anatomical-first approach in the PROMISE (PROspective Multicentre Imaging Study for Evaluation of Chest Pain) trial revealed no significant differences in subsequent cardiovascular events.² However, *post hoc* analyses have since suggested differential benefits associated with anatomical testing among broad subgroups, such as women and patients with diabetes.^{7–9} With both approaches sharing a Class I recommendation,^{1,3} and the increasing availability of CCTA,¹⁰ a data-driven approach to test selection could supplement traditional clinical assessment and maximize the value of current resources.

Causal inference in observational studies is limited by confounding by indication, thus precluding an assessment of how upfront diagnostic decisions may impact long-term outcomes. To address this, Anatomical

vs. Stress teSting decision Support Tool (ASSIST) used a machine-learning-based phenomapping approach to estimate individualized cardiovascular benefit estimates for varying patient profiles in PROMISE benefiting from their random assignment to anatomical vs. functional testing.¹¹ In both PROMISE and a subset of the SCOT-HEART (Scottish Computed Tomography of the Heart) trial, randomization to the strategy recommended by ASSIST as opposed to a different testing modality was associated with a lower risk of major adverse cardiovascular events.¹¹ However, in real-world clinical practice, the choice between functional and anatomical testing is determined in part by the clinician based on their judgement, expected test performance characteristics, as well as local experience and familiarity with each modality.^{12–15} By deploying tools like ASSIST to the real-world, we can describe potential biases and disparities in chest pain management, contrasting real-world practices with those associated with the best outcomes in the controlled environment of an RCT.^{16,17} These findings may support a potential synergy between RCT-derived precision tools and standard clinical pathways.

In the present study, we hypothesize that ASSIST-aligned care would be associated with better clinical outcomes among individuals undergoing non-invasive cardiac testing. To this end, we first describe the agreement between real-world testing practices and the algorithmic approach suggested by ASSIST, its predictors, as well as the association with long-term cardiovascular outcomes across two large and geographically distinct observational cohorts drawn from a US-based hospital network and the UK Biobank (UKB). In an attempt to explore potential mechanisms for outcome differences arising from care practices following ASSIST-aligned care, we further explore its association with rates of downstream CAD diagnosis among patients randomized to anatomical testing who had protocolized follow-up in the setting of PROMISE.²

Methods

Data source and patient population

The overall study approach is outlined in [Figure 1](#). Briefly, the primary objective was to assess the alignment between real-world practices relative to ASSIST-informed algorithmic care, and the downstream association of ASSIST-aligned testing with clinical outcomes. In a mechanistic substudy, we further harnessed the protocolized and complete follow-up of individuals in the RCT environment of PROMISE to explore associations between ASSIST-aligned management and the diagnostic yield of anatomical testing.

Practice evaluation across two multinational cohorts

Yale-New Haven health system

The Yale-New Haven Health electronic health record (EHR; Yale health system) covering five hospitals in Connecticut and Rhode Island was searched for individuals aged 18 years and older who underwent either anatomical (CCTA) or functional testing (treadmill exercise test, stress myocardial perfusion imaging by single-photon emission computed tomography or positron emission tomography, stress echocardiography, or stress cardiac magnetic resonance imaging; either with exercise or pharmacologic stimulation) between 2013 and July 2023. Throughout the study, the first reported test was used to define the diagnostic strategy (anatomical vs. functional), allowing patients to cross-over during their subsequent care as per the PROMISE trial protocol.² To maximize the generalizability of our observations, we included all patients independent of prior history of ischaemic heart disease from both inpatient and outpatient encounters, but performed subgroup analyses across these key subgroups (i.e. outpatient encounters, patients without known ischaemic heart disease, as per PROMISE).² Baseline clinical features were adjudicated based on the last available information prior to the initial cardiac test, as described in [Supplementary material online, Methods](#) and [Table S1](#). Eligible participants were followed for the primary clinical outcome of all-cause mortality or acute myocardial infarction (AMI) following their initial test.

UK Biobank

The UK Biobank is a prospective observational study of over 500 000 people aged 40–69 years recruited in 2006–10 across the UK,¹⁸ offering access to a large cohort of individuals in the community undergoing prospective and protocolized follow-up along with multimodal phenotyping. The data are also linked with longitudinal EHR data for UKB participants. We identified UKB participants who, following their enrolment in the study, underwent non-invasive cardiac testing by functional or anatomical testing as a part of their clinical care (2006–20), which was defined using available inpatient procedure codes until 2021 (see [supplementary material](#)). All baseline characteristics were recorded at the time of enrolment before the performance of the diagnostic test. All eligible participants were followed for the primary clinical outcome of all-cause mortality or AMI. Since the primary cause of death was available in this cohort, we also report a secondary outcome of cardiovascular-specific mortality.

Mechanistic evaluation of diagnostic yield and care in a protocolized setting in PROMISE

We obtained participant-level data from the PROMISE trial (ClinicalTrials.gov identifier: NCT01174550) through the Biologic Specimen and Data Repository Information Coordinating Center of the National Heart, Lung and Blood Institute (NHLBI BioLINCC). Briefly, PROMISE recruited 10 003 patients with suspected CAD who were randomized to either anatomical (CCTA) or functional testing.² We focused on participants who were randomized to the anatomical testing arm and successfully underwent CCTA, thus enabling a detailed assessment of the extent and burden of CAD, an important determinant of downstream cardiovascular outcomes.¹⁹ We collected relevant baseline information (age, sex, demographics, comorbidities, medications, and laboratory measurements) and anatomical information on the severity and extent of epicardial CAD on CCTA.

Agreement between ASSIST prediction and initial test selection

Across all cohorts, we calculated the ASSIST-derived recommendation based on features available before functional testing or CCTA.¹¹ These features include the patient's age at the time of the test, sex, body mass index, history of hypertension, diabetes mellitus, smoking status (active or former), medication use (anti-platelet, statin, or beta-blocker therapy) as well as lipid panel (total cholesterol and HDL levels). The previously trained and validated extreme gradient boosting algorithm (available at: <https://www.cards-lab.org/assist>) computes an individualized effect estimate for the primary endpoint of PROMISE, assuming assignment to anatomical vs. functional testing. Individualized point estimates (hazard ratios, HRs) of <1, favour an anatomical-first approach, whereas HR > 1 favour a functional-first approach. We grouped individuals into those who underwent ASSIST-recommended testing vs. those who did not, depending on whether ASSIST matched the initial diagnostic test performed.

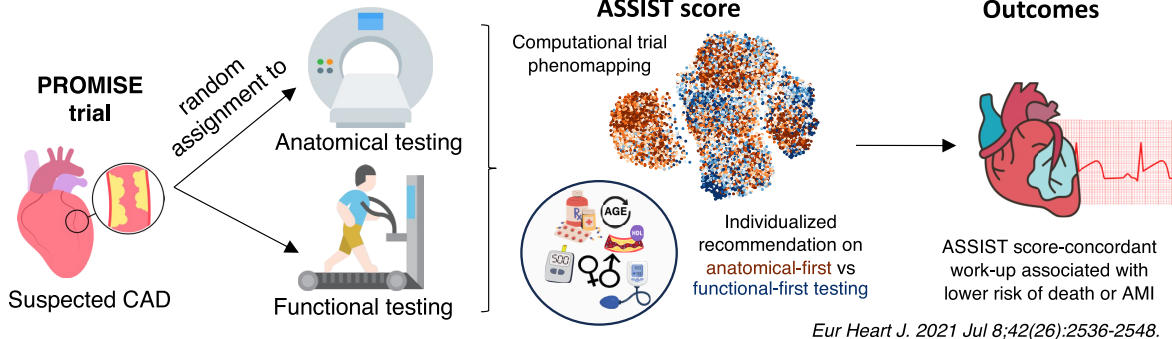
Definition of clinical outcomes

In the Yale health system/UKB analysis, the primary clinical outcome was defined as a composite of time-to-all-cause mortality or AMI, reflecting the prior validation of the ASSIST score.¹¹ The availability of cause of death information in UKB allowed us to also define a secondary outcome of cardiovascular-specific mortality. In the mechanistic substudy, we assessed the association of ASSIST-aligned management with the presence of CAD (defined as any epicardial coronary stenosis 50% or greater on CCTA) among PROMISE participants undergoing protocolized anatomical testing. We also specifically evaluated the presence of stenosis in the left main or proximal left anterior descending coronary artery (LM/pLAD) or LM involvement and/or CAD with three-vessel involvement.

Statistical analysis

Continuous variables are summarized as mean (standard deviation) and categorical variables as counts (valid percentages), with standardized mean difference used to describe the distance between the anatomical and

Background: the ASSIST score



Clinical alignment of real-world testing practices with ASSIST

Cohort #1: US-based health system

5 hospitals (Yale)
134,216 individuals
Age 61 ± 14 years
65,359 (48.7%) female

Cohort #2: UK Biobank

3,901 individuals
Age 66 ± 8 years
1,883 (48.3%) female

Test performed?



ASSIST recommendation?

A digital twin in PROMISE would have predicted improved outcomes using: vs

Referral for ASSIST-recommended testing

Referral for a test other than the one recommended by ASSIST

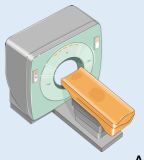
Follow-up for:

Death OR AMI

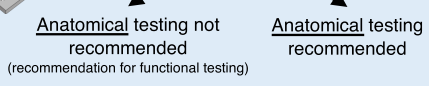
Mechanistic sub-study: Evaluation of ASSIST-aligned care in a protocolized setting

PROMISE trial

Participants with suspected CAD randomized to anatomical assessment by coronary computed tomography angiography



Anatomical versus Stress testing decision Support Tool (ASSIST)



Is ASSIST independently associated with the diagnostic yield of anatomical imaging for CAD (including multivessel or proximal disease)?

Figure 1 Study design overview. AMI, acute myocardial infarction; ASSIST, Anatomical vs. Stress teSting decision Support Tool; CAD, coronary artery disease; EHR, electronic health record; IQR, interquartile range; LAD, left anterior descending artery; PROMISE, PROspective Multicentre Imaging Study for Evaluation of Chest Pain.

functional testing groups across each variable. Categorical features across groups were compared by the χ^2 test, whereas for continuous variable comparisons across two groups, we used Student's t-test or the non-parametric Mann-Whitney test (if deviating from a normal distribution; e.g. pooled cohort equation 10-year risk of atherosclerotic cardiovascular disease risk [PCE ASCVD]).²⁰ The correlation between PCE ASCVD and continuous ASSIST predictions was assessed using Spearman's rho. Missing categorical and continuous values were imputed using chained equation imputation with random forests before inclusion in multivariable regression models.²¹

Within each cohort (Yale health system and UKB), we calculated a propensity score (probability) for undergoing anatomical vs. functional-first testing by fitting a multivariable logistic regression model adjusted for all ASSIST components (age at the time of the test, sex, body mass index, history of hypertension, diabetes mellitus, smoking status, anti-platelet, statin

or beta-blocker use, total cholesterol, and HDL levels) as well as race and ethnic background, ischaemic heart disease history, and chronic kidney disease (see [supplementary material](#) for further details).^{22,23} We report McFadden's pseudo- R^2 as a marker of goodness-of-fit for the model, further adjusting for the study year. Next, to explore predictors of undergoing ASSIST-recommended testing, we fit a multivariable logistic regression model using ASSIST-consistent testing as the dependent variable and the previously mentioned predictors as independent variables. The association of ASSIST-consistent management with the incidence of all-cause mortality or AMI was assessed in multivariable Cox regression models, both unadjusted and adjusted for the above covariates and propensity score for anatomical testing. The proportional hazards assumption was assessed using Schoenfeld residuals at the rounded median follow-up of 5 years with a P-value of >0.05, suggesting that the proportional hazard assumption was met for both cohorts (P = 0.11 and 0.52, respectively). Adjusted cumulative

hazard curves are graphically presented. The association between ASSIST-consistent testing and the primary outcome was further assessed across clinically relevant subgroups, with results summarized in the form of forest plots and *P*-values for heterogeneity derived from the corresponding (ASSIST-alignment \times subgroup) interaction term. This was restricted to the Yale health system analysis, given the higher counts of events.

In the mechanistic substudy analysis in PROMISE, the association between undergoing the test recommended by ASSIST (vs. a different strategy) with the presence of CAD was assessed in logistic regression models adjusted for the same covariates (with glomerular filtration rate replacing the chronic kidney disease flag) but no propensity score adjustment given that all patients underwent CCTA. We also present unadjusted proportions for individuals meeting these definitions in each ASSIST group with 95% binomial confidence intervals (CIs).

All statistical tests were two-sided with a significance level of 0.05. Analyses were performed using Python (version 3.11.2) and R (version 4.2.3). The analyses stand consistent with the STROBE and CODE-EHR statements.^{24,25}

Consent

The Yale Institutional Review Board approved our study and waived the requirement for informed consent for our *post hoc* analysis of de-identified data for all arms of this study. The UKB analysis was conducted under research application #71033.

Results

Anatomical vs. functional testing in a real-world setting

The Yale health system cohort included 134 216 patients [mean age 61 ± 14 years, $n = 65\,359$ (48.7%) female] who underwent functional or anatomical assessment between 2013 and 14 July 2023 [$n = 28\,442$ (21.2%) as inpatient]. Most diagnostic tests were functional [$n = 130\,196$ (97.0%)], among which 59 825 (46.0%) were nuclear myocardial perfusion imaging studies (Table 1). The proportion of anatomical testing by CCTA ($n = 4\,020$, 3.0%) increased over time from 186 out of 19 458 (1.0%) in 2013–14 to 763 out of 7460 (10.2%) in 2022–23 (see Supplementary material online, Figure S1). Compared with the functional test arm, individuals in the anatomical arm were younger (59 ± 16 vs. 61 ± 14 years, $P < 0.001$), with lower prevalence of a baseline diagnosis of hypertension [1214 (30.2%) vs. 50 178 (38.5%)], diabetes mellitus [306 (7.6%) vs. 13 755 (10.6%)], statin [1184 (29.5%) vs. 51 389 (39.5%)], or anti-platelet therapy at baseline [629 (15.6%) vs. 30 054 (23.1%); all $P < 0.001$]. Furthermore, among patients without a history of ischaemic heart disease, the median imputed 10-year PCE ASCVD risk was 4.2% (25th–75th percentile: 1.6–9.4%) and 5.0% (2.1–10.4%), respectively ($P < 0.001$). In the UKB cohort, among 3901 individual patients, 581 (14.9%) underwent anatomical and 3320 (85.1%) functional testing, with a mean age of 66 ± 8 years in each group and 262 (45.1%) vs. 1621 (48.8%) females, respectively (Table 1). Clinical factors only explained 1.5% and 1.3% of the variance in testing selection, increasing to 10.4% and 3.9% after incorporating the testing year in the Yale and UKB cohorts, respectively.

Association between ASSIST-recommended and real-world testing strategies

In the Yale health system cohort, the PROMISE-derived ASSIST framework would have projected better outcomes with anatomical-first testing in 23 859 (17.8%) individuals, nearly 6 times as many patients as the ones who were referred for CCTA ($n = 4020$), and functional-first testing in 110 357 (82.2%) individuals. In the UKB, ASSIST would have

favoured anatomical testing in 827 (21.2%) individuals, a number 1.4-fold higher than the actual number of CCTA-first testing ($n = 581$), vs. functional testing in the remaining 3074 (78.8%) of the included participants. There was a moderate correlation between ASSIST-derived individualized heart ratios favouring anatomical testing and higher cardiovascular risk, as assessed by the PCE-derived 10-year ASCVD risk (Spearman's ρ 0.46, $P < 0.001$).

In summary, 107 573 (80.1%) and 2715 (70.0%) individuals had the testing strategy that was recommended by ASSIST in the Yale health system and UKB, respectively. In multivariable regression analyses further adjusted for the propensity of undergoing anatomical vs. functional testing, female sex and a history of diabetes were all significantly associated with a lower likelihood of receiving the test recommended by ASSIST across both cohorts. Furthermore, in the Yale health system cohort, younger patients as well as Black individuals were less likely to undergo the ASSIST-recommended testing strategy (Table 2).

Alignment with ASSIST-based care and long-term outcomes

In the Yale health system cohort, among 134 216 individuals followed over a median of 4.9 [interquartile range (IQR): 2.4–7.1] years, 11 392 participants experienced the primary outcome, which included 3788 confirmed AMI events and 8825 deaths. In the UKB cohort, median follow-up was 5.4 (IQR: 2.6–8.8) years, with 478 primary outcome events (including 131 AMI events, and 385 deaths, 148 of which were attributed to a primary cardiovascular aetiology) among 3901 individuals. An ASSIST profile favouring functional (vs. anatomical) testing and assignment to functional (vs. anatomical testing) were all independently associated with a lower risk of the primary outcome [HR (95% CI) of 0.82 (0.78–0.86), $P < 0.001$ and 0.62 (0.52–0.70), $P < 0.001$ in the Yale health system, and 0.55 (0.44–0.71), $P < 0.001$, and 0.73 (0.58–0.91), $P < 0.001$, in the UKB, respectively; see Supplementary material online, Figure S2].

Overall, in both the Yale health system and UKB cohorts, patients who underwent testing consistent with the ASSIST recommendation were less likely to experience the primary outcome of all-cause mortality or AMI [HR_{adjusted} of 0.81 (95% CI 0.77–0.85), $P < 0.001$ and 0.74 (95% CI 0.60–0.90), $P = 0.003$, respectively] after propensity score adjustment that also included all individual covariates (Figure 2 and Supplementary material online, Table S2). This association persisted for a secondary outcome of cardiovascular-specific mortality in the UKB [HR_{adjusted} of 0.55 (95% CI 0.39–0.78), $P < 0.001$].

Subgroup analyses

We further explored the association between ASSIST-aligned vs. discordant management with the primary outcome across clinically relevant subgroups in the Yale health system (Figure 3). We observed no significant heterogeneity across sex groups ($p_{\text{interaction}} = 0.15$), diabetes mellitus history ($p_{\text{interaction}} = 0.63$), year the study was performed ($p_{\text{interaction}} = 0.10$), type of functional modality performed ($p_{\text{interaction}} = 0.71$), inclusive of exercise or pharmacological stress testing ($p_{\text{interaction}} = 0.58$). However, the association between ASSIST-aligned testing and lower risk of the primary outcome was more pronounced among patients undergoing inpatient [HR_{adjusted} 0.71 (95% CI 0.66–0.76)] vs. outpatient testing [0.88 (95% CI 0.83–0.94)], $p_{\text{interaction}} < 0.001$, with a graded relationship across PCE-derived ASCVD risk profiles [0.60 (0.54–0.66) for <5% vs. 0.91 (0.84–0.99), $p_{\text{interaction}} < 0.001$; Supplementary material online, Figure S3].

Table 1 Baseline characteristics for the Yale health system and UK Biobank cohorts

	Yale health system cohort			UKB cohort		
	Anatomical	Functional	SMD	Anatomical	Functional	SMD
Total counts	4020	130 196	—	581	3320	—
Type of functional test			—			
Stress echo	0 (0.0)	33 537 (25.8)		0 (0.0)	1657 (49.9)	—
Nuclear MPI	0	59 825 (46.0)		0 (0.0)	1572 (47.3)	
Stress CMR	0 (0.0)	265 (0.2)		—	—	
Exercise treadmill test	0 (0.0)	36 569 (28.1)		0 (0.0)	91 (2.7)	
Age (years)	58.5 (16.1)	61.0 (14.0)	0.165	66.2 (8.3)	65.9 (7.8)	0.045
Female (sex)	1853 (46.1)	63 506 (48.8)	0.054	262 (45.1)	1621 (48.8)	0.075
Hispanic ethnicity	447 (11.1)	12 687 (9.7)	0.054	—	—	—
Race/ethnic background			0.058			0.082
White	2991 (74.4)	100 015 (76.8)		548 (94.3)	3101 (93.4)	
Asian	96 (2.4)	3055 (2.3)		15 (2.6)	95 (2.9)	
Black	477 (11.9)	13 750 (10.6)		10 (1.7)	45 (1.4)	
Other/unknown	456 (11.3)	13 376 (10.3)		7 (1.2)	71 (2.1)	
Mixed				1 (0.2)	8 (0.2)	
Hypertension	1214 (30.2)	50 178 (38.5)	0.176	298 (51.3)	1472 (44.3)	
Diabetes mellitus	306 (7.6)	13 755 (10.6)	0.103	95 (16.4)	489 (14.7)	0.045
Prior ischaemic heart disease (incl. PCI/CABG)	184 (4.6)	8316 (6.4)	0.08	250 (43.0)	1380 (41.6)	0.03
Prior AMI	43 (1.1)	1215 (0.9)	0.014	9 (1.5)	19 (0.6)	0.095
Stroke	174 (4.3)	3823 (2.9)	0.074	52 (9.0)	122 (3.7)	0.218
CKD	58 (1.4)	2612 (2.0)	0.043	47 (8.1)	183 (5.5)	0.103
PAD	3 (0.1)	54 (0.0)	0.014	14 (2.4)	47 (1.4)	0.073
Active smoking	8 (0.2)	422 (0.3)	0.024	70 (12.0)	412 (12.4)	0.011
Former smoking	124 (3.1)	3919 (3.0)	0.004	253 (43.5)	1398 (42.1)	0.029
Total cholesterol (mg/dL)	177 (47)	179 (44)	0.049	209 (47)	211 (50)	0.032
HDL (mg/dL)	53 (19)	55 (19)	0.086	52 (15)	52 (14)	0.007
BMI (kg/m ²)	29.3 (6.6)	30.0 (7.0)	0.102	28.4 (5.0)	28.8 (5.2)	0.083
Beta-blockers	1245 (31.0)	36 994 (28.4)	0.056	103 (17.7)	689 (20.8)	0.077
Statin use	1184 (29.5)	51 389 (39.5)	0.212	214 (36.8)	1353 (40.8)	0.081
Anti-platelet use	629 (15.6)	30 054 (23.1)	0.189	187 (32.2)	1219 (36.7)	0.095

Variables are summarized as mean (standard deviation) or counts (valid percentages, %) as appropriate.

AMI, acute myocardial infarction; BMI, body mass index; CABG, coronary artery bypass grafting; CKD, chronic kidney disease; MPI, myocardial perfusion imaging; PAD, peripheral arterial disease; PCI, percutaneous coronary intervention; SMD, standardized mean difference; UKB, UK Biobank.

ASSIST-aligned care and diagnostic yield of protocolized anatomical testing

Building on prior evidence demonstrating an association between the anatomical severity and burden of coronary atherosclerosis with adverse clinical outcomes,^{19,26} we further investigated the association between the ASSIST recommendation and the diagnostic yield of protocolized CCTA imaging in the PROMISE study. In 4734 patients who underwent CCTA imaging in the anatomical arm of PROMISE [median age 59 (54–65) years, $n = 2426$ (51.2%) female; see [Supplementary material online, Table S3](#)], 1329 (28.1%) patients had evidence of at least one epicardial coronary stenosis 50% or greater, 510 (10.8%) patients had disease involving the left main or proximal LAD, and 96 (2.0%) left main and/or 3-vessel disease involvement. Patients for whom ASSIST would have recommended anatomical ($n = 2005$, 42.4%) vs. functional testing ($n = 2729$, 57.6%) were 17% more likely to have any CAD [OR_{adjusted} 1.17 (95% CI 1.02–1.35), $P = 0.023$], 30% more likely to have disease in the left main or proximal LAD [OR_{adjusted} 1.30 (95% CI 1.07–1.58), $P = 0.009$], and (numerically)

38% more likely to have left main or 3-vessel CAD [OR_{adjusted}:1.38 (95% CI 0.90–2.12), $P = 0.13$], although this did not reach statistical significance ([Figure 4](#)).

Discussion

In this descriptive analysis of more than 130 000 patients with suspected CAD who predominantly underwent functional testing across two geographically divergent cohorts, we demonstrate reproducible disparities in clinical outcomes when the initial diagnostic test deviated from the one suggested by ASSIST, an RCT-derived algorithm that links personalized clinical profiles to the diagnostic strategy associated with the best clinical outcomes. By applying this framework, we show a dissociation between algorithmic and real-world testing practices that is more pronounced among patients who were younger, female, and of Black race, suggesting that the slow uptake of CCTA may have disproportionately disadvantaged select demographic groups. Analyses across distinct geographical, temporal, and clinical settings demonstrate

Table 2 Predictors of undergoing Anatomical vs. Stress teSting decision Support Tool--recommended (vs. discordant) testing

Predictor	Yale health system	UK Biobank cohort
	Beta (SE), OR (95% CI), P-value	Beta (SE), OR (95% CI), P-value
Age (per 10 years)	0.14 (0.01), 1.15 (1.13–1.17), $P < 0.001$	0.05 (0.05), 1.05 (0.95–1.16), $P = 0.34$
Female (sex)	–1.64 (0.02), 0.19 (0.19–0.20), $P < 0.001$	–0.49 (0.08), 0.61 (0.52–0.72), $P < 0.001$
Hispanic ethnicity	0.03 (0.04), 1.03 (0.96–1.11), $P = 0.392$	—
Race: Asian (vs. White)	–0.12 (0.06), 0.89 (0.80–0.99), $P = 0.038$	0.35 (0.23), 1.42 (0.91–2.26), $P = 0.13$
Race: Black (vs. White)	–0.23 (0.03), 0.79 (0.75–0.83), $P < 0.001$	–0.01 (0.32), 0.99 (0.54–1.88), $P = 0.98$
BMI (per 10 kg/m ²)	0.27 (0.02), 1.31 (1.27–1.35), $P < 0.001$	1.13 (1.07), 3.09 (0.55–57.91), $P = 0.29$
Total cholesterol (per 10 mg/dL)	–0.02 (0.00), 0.98 (0.97–0.98), $P < 0.001$	0.02 (0.01), 1.02 (1.00–1.04), $P = 0.12$
HDL (per 10 mg/dL)	0.03 (0.01), 1.03 (1.01–1.04), $P < 0.001$	0.06 (0.04), 1.06 (0.99–1.14), $P = 0.08$
Hypertension	–1.16 (0.02), 0.31 (0.30–0.33), $P < 0.001$	–0.09 (0.09), 0.92 (0.78–1.09), $P = 0.32$
Diabetes mellitus	–2.34 (0.02), 0.10 (0.09–0.10), $P < 0.001$	–1.3 (0.11), 0.27 (0.22–0.34), $P < 0.001$
Ischaemic heart disease history	0.40 (0.03), 1.49 (1.40–1.59), $P < 0.001$	0.00 (0.09), 1.00 (0.85–1.19), $P = 0.97$
Chronic kidney disease	0.19 (0.05), 1.21 (1.09–1.34), $P < 0.001$	–0.21 (0.15), 0.81 (0.60–1.10), $P = 0.17$
Anti-platelet use	–0.47 (0.02), 0.63 (0.60–0.65), $P < 0.001$	–0.42 (0.10), 0.66 (0.55–0.79), $P < 0.001$
Statin use	–0.19 (0.02), 0.83 (0.80–0.86), $P < 0.001$	0.38 (0.10), 1.46 (1.20–1.79), $P < 0.001$
Beta-blocker use	–1.91 (0.02), 0.15 (0.14–0.15), $P < 0.001$	–1.06 (0.10), 0.35 (0.29–0.42), $P < 0.001$
Active smoking	–0.62 (0.12), 0.54 (0.42–0.69), $P < 0.001$	–0.29 (0.12), 0.75 (0.59–0.95), $P = 0.02$
Former smoking	–0.38 (0.04), 0.68 (0.63–0.74), $P < 0.001$	–0.24 (0.08), 0.78 (0.67–0.92), $P = 0.003$

BMI, body mass index; CI, confidence interval; OR, odds ratio; SE, standard error.

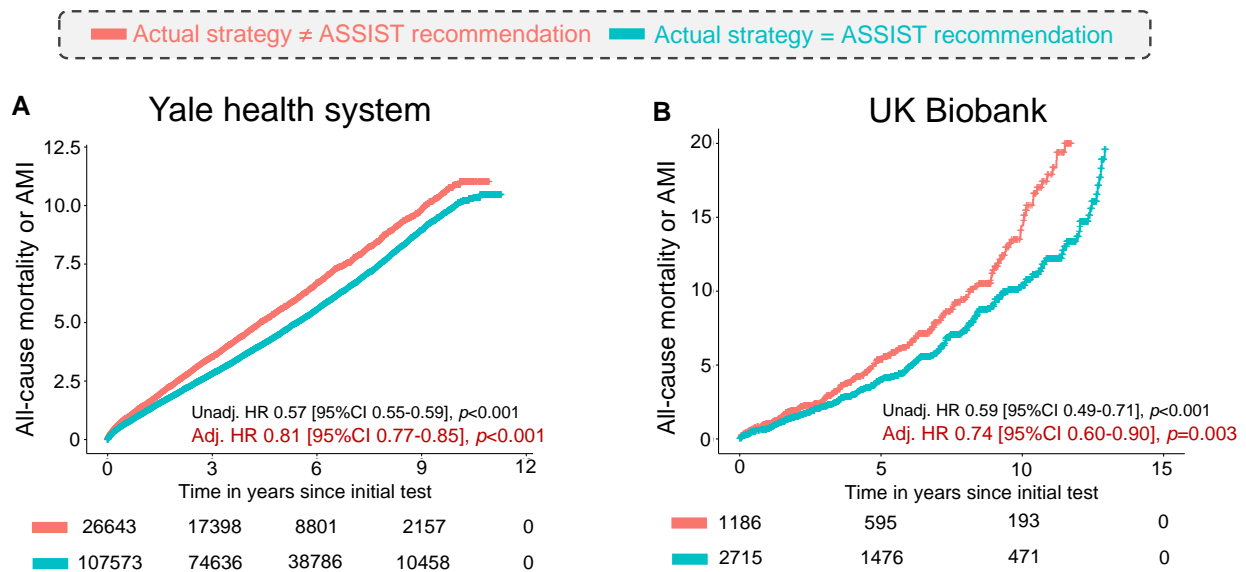


Figure 2 Association between Anatomical vs. Stress teSting decision Support Tool-aligned testing and adverse clinical outcomes. Adjusted Cox regression-derived cumulative hazard curves for all-cause mortality or acute myocardial infarction in the (A) Yale health system, and (B) a subset of the UK Biobank undergoing non-invasive cardiac testing. ASSIST, Anatomical vs. Stress teSting decision Support Tool; CI, confidence interval; HR, hazard ratio.

consistently worse clinical outcomes, not explained by traditional cardiovascular factors. In mechanistic analyses of patients undergoing anatomical testing with protocolized follow-up in the setting of an RCT, we further show that ASSIST-alignment is associated with greater

diagnostic yield for CAD or LM/pLAD disease. These findings provide a paradigm on how to evaluate clinical care pathways using inference drawn from high-quality RCTs, and an insight into how data-driven decision-making might supplement the traditional clinical evaluation.

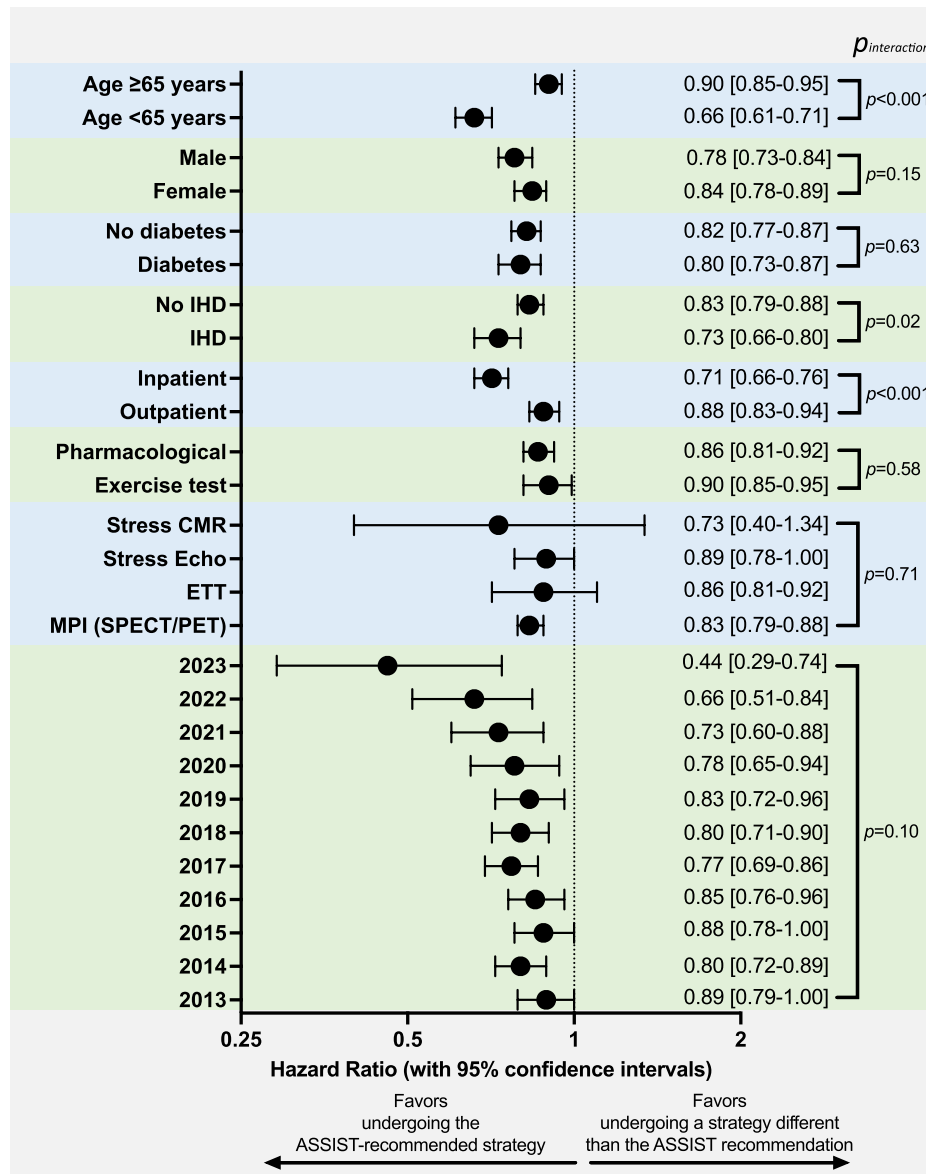


Figure 3 Subgroup analysis in the Yale health system cohort. Forest plot showing the adjusted hazard ratios and corresponding 95% confidence intervals for the association between Anatomical vs. Stress teSting decision Support Tool-aligned testing and all-cause mortality or acute myocardial infarction across relevant subgroups. A P -value for heterogeneity (derived from an interaction test) is also presented. ASSIST, Anatomical vs. Stress teSting decision Support Tool; CI, confidence interval; ETT, exercise treadmill test; HR, hazard ratio; IHD, ischaemic heart disease; MPI, myocardial perfusion imaging; PET, positron emission tomography; SPECT, single-photon emission computed tomography.

Taken together, our findings highlight how routine real-world practices might differ from patterns associated with best outcomes within the controlled and protocolized environment of RCTs, and how this may contribute to differences in downstream outcomes, as consistently observed across RCTs and diverse cohorts.²⁷ When studying the downstream effects of any diagnostic test,²⁸ disparities in subsequent outcomes are inherently related to the complex bundle of care that follows, rather than the test itself.²⁹ Such factors include patient education, risk factor and lifestyle modification, medication adjustments, as well as further testing and procedures.^{30,31} Furthermore, real-world practice patterns reflect multiple patient-, site-, and provider-specific parameters that collectively determine the propensity to refer for anatomical vs. functional testing.^{32–34} ASSIST enables the direct

translation of heterogeneous treatment effects noted in subpopulations enrolled in an RCT, where the use of these tests is independent of patient characteristics and possible provider bias. Building on our prior work on phenomapping-guided inference from RCTs,^{11,35–37} ASSIST allows us to project each new participant to a simulated PROMISE environment and predict the strategy that would work best in this simulated setting. This generates a mechanism to emulate real-world outcomes relative to an environment, where the key clinical decision is made independent of clinical bias or knowledge and is guided by the diagnostic or therapeutic effects on individuals.¹¹ This can then be used to flag, quantify, and possibly assess human-led decision-making against an algorithmic approach to test selection.^{38–41} Our findings are supported by *post hoc* analyses of PROMISE, showing that an objective

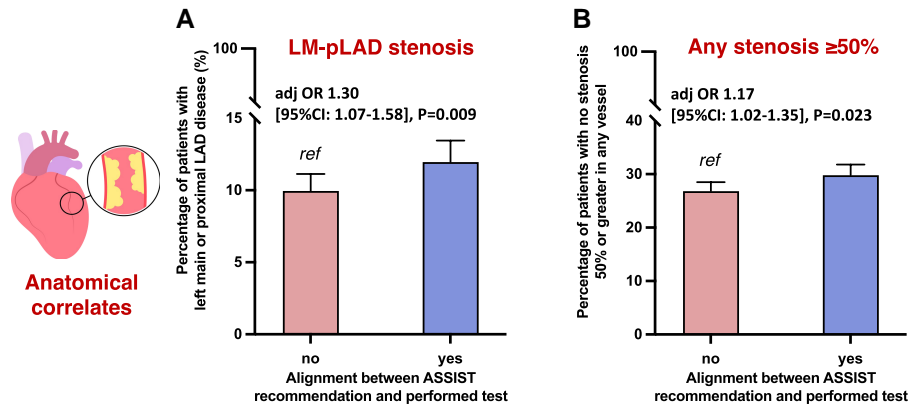


Figure 4 Mechanistic evaluation of Anatomical vs. Stress testing decision Support Tool with anatomical yield in the PROMISE (Prospective Multicentre Imaging Study for Evaluation of Chest Pain) trial. Association between ASSIST-agreement with anatomical evaluation and (A) the presence of coronary stenosis in the left main or proximal left anterior descending artery, or (B) any significant epicardial coronary artery disease. Bar graphs denote proportions for individuals meeting these definitions in each ASSIST group with 95% binomial confidence intervals. ASSIST, Anatomical vs. Stress testing decision Support Tool; CAD, coronary artery disease; CI, confidence interval; LM, left main artery; pLAD, proximal left anterior descending artery; OR, odds ratio.

rather than subjective approach to diagnostic test selection is associated with a higher rate of CAD detection that may modify subsequent behaviour and management.

The broader clinical implications of this work are two-fold. First, our work suggests potential value in accounting for the ASSIST phenotype when choosing between anatomical and functional testing, offering a data-driven, personalized inference approach that can interpret landmark trial data in an individualized manner for each patient. While the current evidence is insufficient to justify using ASSIST as a decision support tool, the study represents data in support of a future nudged-based, pragmatic RCT that compares ASSIST-informed test allocation vs. routine care among cohorts that have traditionally favoured a functional-first approach. Second, our approach demonstrates specific patient subgroups in whom established clinical practices may perpetuate outcome disparities.^{38–41} As shown in our analysis, the association between ASSIST-aligned testing and outcomes is consistent across such demographic groups, yet female patients or patients with diabetes are less likely to undergo the testing strategy recommended by ASSIST, compared with their counterparts. In fact, our framework is unique in that it links these to potentially modifiable management decisions (as studied in RCTs), thus suggesting a data-driven way to improve the quality of care. This can be particularly challenging to study in the context of observational registries, given that most studies are either cross-sectional or focus on the short-term prevalence of obstructive CAD,^{1,3,42} rather than the long-term incidence of major clinical events. The focus on obstructive CAD may misrepresent the modifiable portion of a patient's cardiovascular risk, with numerous studies showing a lack of mortality benefit with revascularization in most stable coronary disease cases.^{43,44} We thus infer that ASSIST may be flagging interactions between a patient's profile and common management strategies that collectively explain differential clinical outcomes.

Our study has certain limitations that merit consideration. First, there are key differences between the cohorts used here and the trial populations in which ASSIST was developed. For instance, both the Yale health cohort and UKB are skewed towards functional testing, likely reflecting local practices, availability, and experience at the time of the study. Both cohorts also included scans after 2010–13 when the PROMISE trial was conducted, thus reflecting the evolution in clinically available technologies. This highlights the need for a longitudinal evaluation of ASSIST to assess

how changes in practice patterns might affect its association with outcomes. Second, to maximize the generalizability of our observations, we included a broader population than one studied in PROMISE² and performed multivariable adjustments, including propensity score adjustments for the probability of undergoing CCTA vs. functional testing. This allowed us to track outcomes both in outpatient populations resembling PROMISE,² but also commonly encountered scenarios, such as the hospital-based management of stable chest pain. Furthermore, with follow-up that was more than twice as long as PROMISE, our analysis enabled us to explore not just the short-term but also the intermediate- and long-term outcome associations of upfront testing strategies. Third, the *post hoc* and retrospective associations described in this study do not necessarily imply a causal mechanism. Indeed, we acknowledge that propensity score adjustment does not address unmeasured sources of confounding, which likely explains most of the variance seen in the use of anatomical vs. functional testing in our cohorts. Moreover, comparing anatomical characterization across subgroups was not done in the Yale health system or UKB due to a lack of consistently reported information and direct confounding by the provider-driven test selection. Therefore, this was performed in the setting of an RCT, where all patients got randomized to CCTA irrespective of their presentation and clinical profile. We do acknowledge, however, that analyses in PROMISE are limited by their *post hoc* nature, the fact that ASSIST was trained in the same cohort (although without incorporating any information about the results of the CCTA scans) and that differences in CAD detection rates may not fully explain the observed disparities in outcomes. Fourth, the cause of death could not be reliably ascertained in the EHR-based cohort; however, cardiovascular-specific mortality was presented as a secondary outcome in the UKB. Finally, patients may have undergone different tests later during their course. However, we only focused on the initial testing strategy to define the testing approach, consistent with the PROMISE trial,² which was leveraged for the development of ASSIST.¹¹

Conclusions

In summary, we demonstrate that misalignment between real-world practices around non-invasive testing for suspected CAD and an RCT-derived algorithmic strategy are associated with worse

cardiovascular outcomes across multinational cohorts. This discrepancy is more pronounced among select clinical and demographic groups, highlighting that subjectivity in selection may perpetuate biases and impact the long-term cardiovascular trajectory of select populations. These hypothesis-generating findings may inform future clinical trials assessing a more equitable and data-driven deployment of cardiac testing.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Funding

The authors acknowledge funding through the National Institutes of Health (NIH) under awards 1F32HL170592-01 (E.K.O.), R01HL167858 (R.K.) and K23HL153775 (R.K.), and the Doris Duke Charitable Foundation under award 2022060 (R.K.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

Conflict of interest: E.K.O. is a co-inventor in patent applications (US17/720,068, 63/619,241, 63/177,117, 63/580,137, 63/606,203, 63/562,335, WO2018078395A1, and WO2020058713A1), has been a consultant for Caristo Diagnostics Ltd and Ensignt-AI Inc, and received royalty fees for technology licensed through the University of Oxford related to computed tomography phenotyping of cardiometabolic risk. E.K.O. and R.K. are academic cofounders of Evidence2Health, a health data analytics company. R.K. is an associate editor of JAMA. He receives research support through Yale, from Bristol-Myers Squibb, Novo Nordisk, and BridgeBio. He is a co-inventor of US Provisional Patent Applications 63/177,117, 63/428,569, 63/346,610, 63/484,426, 63/508,315, 63/606,203, 63/619,241, 63/562,335 and co-founder of Ensignt-AI, Inc. H.M.K. works under contract with the Centers for Medicare & Medicaid Services to support quality measurement programmes, was a recipient of a research grant from Johnson & Johnson, through Yale University, to support clinical trial data sharing; was a recipient of a research agreement, through Yale University, from the Shenzhen Center for Health Information for work to advance intelligent disease prevention and health promotion; collaborates with the National Center for Cardiovascular Diseases in Beijing; receives payment from the Arnold & Porter Law Firm for work related to the Sanofi clopidogrel litigation, from the Martin Baughman Law Firm for work related to the Cook Celect IVC filter litigation, and from the Siegfried and Jensen Law Firm for work related to Vioxx litigation; chairs a Cardiac Scientific Advisory Board for UnitedHealth; was a member of the IBM Watson Health Life Sciences Board; is a member of the Advisory Board for Element Science, the Advisory Board for Facebook, and the Physician Advisory Board for Aetna; and is the co-founder of Hugo Health, a personal health information platform, and co-founder of Refactor Health, a healthcare AI-augmented data management company, and Ensignt-AI, Inc. All other authors declare no relevant competing interests.

Data availability

The PROMISE data are available through the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC) of the National Heart, Lung, and Blood Institute (NHLBI). The Yale health system data cannot be made available since they contain personal identifiable information. The UKB data are publicly available for research use. The code for the statistical analysis can be found at: https://github.com/CarDS-Yale/ASSIST_Real_World_exploration.

References

- Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK, et al. 2021 AHA/ACC/AASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. *Circulation* 2021;**144**: e368–e454.
- Douglas PS, Hoffmann U, Patel MR, Mark DB, Al-Khalidi HR, Cavanaugh B, et al. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med* 2015;**372**:1291–1300.
- Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J* 2020;**41**:407–477.
- Jørgensen ME, Andersson C, Nørgaard BL, Abdulla J, Shreibati JB, Torp-Pedersen C, et al. Functional testing or coronary computed tomography angiography in patients with stable coronary artery disease. *J Am Coll Cardiol* 2017;**69**:1761–1770.
- Karády J, Mayrhofer T, Ivanov A, Foldyna B, Lu MT, Ferencik M, et al. Cost-effectiveness analysis of anatomic vs functional Index testing in patients with low-risk stable chest pain. *JAMA Netw Open* 2020;**3**:e2028312.
- Celeng C, Leiner T, Maurovich-Horvat P, Merykly B, de Jong P, Dankbaar JW, et al. Anatomical and functional computed tomography for diagnosing hemodynamically significant coronary artery disease: a meta-analysis. *JACC Cardiovasc Imaging* 2019;**12**: 1316–1325.
- Pagidipati NJ, Hemal K, Coles A, Mark DB, Dolor RJ, Pellikka PA, et al. Sex differences in functional and CT angiography testing in patients with suspected coronary artery disease. *J Am Coll Cardiol* 2016;**67**:2607–2616.
- Hemal K, Pagidipati NJ, Coles A, Dolor RJ, Mark DB, Pellikka PA, et al. Sex differences in demographics, risk factors, presentation, and noninvasive testing in stable outpatients with suspected coronary artery disease: insights from the PROMISE trial. *JACC Cardiovasc Imaging* 2016;**9**:337–346.
- Sharma A, Coles A, Sekaran NK, Pagidipati NJ, Lu MT, Mark DB, et al. Stress testing versus CT angiography in patients with diabetes and suspected coronary artery disease. *J Am Coll Cardiol* 2019;**73**:893–902.
- Foldyna B, Uhlig J, Mayrhofer T, Natale L, Vliegthart R, Lotz J, et al. Rising utilization of coronary CT angiography across Europe over the last decade: insights from a large prospective European registry. *Eur Heart J* 2021;**42**:203.
- Oikonomou EK, Van Dijk D, Parise H, Suchard MA, de Lemos J, Antoniadis C, et al. A phenomapping-derived tool to personalize the selection of anatomical vs. Functional testing in evaluating chest pain (ASSIST). *Eur Heart J* 2021;**42**:2536–2548.
- Hendel Robert C, Berman Daniel S, Di Carli Marcelo F, Heidenreich Paul A, Henkin Robert E, Pellikka Patricia A, et al. ACCF/ASNC/ACR/AHA/AASE/SCCT/SCMR/SNM 2009 appropriate use criteria for cardiac radionuclide imaging. *J Am Coll Cardiol* 2009; **53**:2201–2229.
- Taylor AJ, Cerqueira M, Hodgson JM, Mark D, Min J, O’Gara P, et al. ACCF/SCCT/ACR/AHA/AASE/ASNC/NASCI/SCAI/SCMR 2010 appropriate use criteria for cardiac computed tomography. *J Cardiovasc Comput Tomogr* 2010;**4**:407.e1–407.e33.
- Wolk MJ, Bailey SR, Doherty JU, Douglas PS, Hendel RC, Kramer CM, et al. ACCF/AHA/AASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2014;**63**:380–406.
- Patel MR, Calhoon JH, Dehmer GJ, Grantham JA, Maddox TM, Maron DJ, et al. ACC/AATS/AHA/AASE/ASNC/SCAI/SCCT/STS 2017 appropriate use criteria for coronary revascularization in patients with stable ischemic heart disease: a report of the American College of Cardiology Appropriate Use Criteria Task Force, American Association for Thoracic Surgery, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, and Society of Thoracic Surgeons. *J Nucl Cardiol* 2017;**24**:1759–1792.
- Chapman EN, Kaatz A, Carnes M. Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. *J Gen Intern Med* 2013;**28**:1504–1510.
- Clerc Liaudat C, Vaucher P, De Francesco T, Jaunin-Stalder N, Herzog L, Verdon F, et al. Sex/gender bias in the management of chest pain in ambulatory care. *Women Health (Lond)* 2018;**14**:1745506518805641.
- Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779.
- Reynolds HR, Shaw LJ, Min JK, Page CB, Berman DS, Chaitman BR, et al. Outcomes in the ISCHEMIA trial based on coronary artery disease and ischemia severity. *Circulation* 2021;**144**:1024–1038.

20. Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min Y-I, Basu S. Clinical implications of revised pooled cohort equations for estimating atherosclerotic cardiovascular disease risk. *Ann Intern Med* 2018;**169**:20–29.
21. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012;**28**:112–118.
22. Yao XI, Wang X, Speicher PJ, Hwang ES, Cheng P, Harpole DH, et al. Reporting and guidelines in propensity score analysis: a systematic review of cancer and cancer surgical studies. *J Natl Cancer Inst* 2017;**109**:djw323.
23. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, et al. Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. *J Am Coll Cardiol* 2017;**69**:345–357.
24. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–808.
25. Kotecha D, Asselbergs FW, Achenbach S, Anker SD, Atar D, Baigent C, et al. CODE-EHR best-practice framework for the use of structured electronic health-care records in clinical research. *Lancet Digit Health* 2022;**4**:e757–e764.
26. Ferencik M, Mayrhofer T, Bittner DO, Emami H, Puchner SB, Lu MT, et al. Use of high-risk coronary atherosclerotic plaque detection for risk stratification of patients with stable chest pain: a secondary analysis of the PROMISE randomized clinical trial. *JAMA Cardiol* 2018;**3**:144–152.
27. Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *N Engl J Med* 2020;**382**:674–678.
28. Rodger M, Ramsay T, Fergusson D. Diagnostic randomized controlled trials: the final frontier. *Trials* 2012;**13**:137.
29. Lavallée JF, Gray TA, Dumville J, Russell W, Cullum N. The effects of care bundles on patient outcomes: a systematic review and meta-analysis. *Implement Sci* 2017;**12**:142.
30. Roifman I, Wijeyesundera HC, Austin PC, Rezaei MR, Wright GA, Tu JV. Comparison of anatomic and clinical outcomes in patients undergoing alternative initial noninvasive testing strategies for the diagnosis of stable coronary artery disease. *J Am Heart Assoc* 2017;**6**:e005462.
31. Ladapo JA, Hoffmann U, Lee KL, Coles A, Huang M, Mark DB, et al. Changes in medical therapy and lifestyle after anatomical or functional testing for coronary artery disease. *J Am Heart Assoc* 2016;**5**:e003807.
32. Hall MK, Kea B, Wang R. Recognising bias in studies of diagnostic tests part 1: patient selection. *Emerg Med J* 2019;**36**:431–434.
33. Kea B, Hall MK, Wang R. Recognising bias in studies of diagnostic tests part 2: interpreting and verifying the index test. *Emerg Med J* 2019;**36**:501–505.
34. Ladapo JA, Blecker S, Elashoff MR, Federspiel JJ, Vieira DL, Sharma G, et al. Clinical implications of referral bias in the diagnostic performance of exercise testing for coronary artery disease. *J Am Heart Assoc* 2013;**2**:e000505.
35. Oikonomou EK, Spatz ES, Suchard MA, Khera R. Individualising intensive systolic blood pressure reduction in hypertension using computational trial phenomaps and machine learning: a post-hoc analysis of randomised clinical trials. *Lancet Digit Health* 2022;**4**:e796–e805.
36. Oikonomou EK, Suchard MA, McGuire DK, Khera R. Phenomapping-derived tool to individualize the effect of canagliflozin on cardiovascular risk in type 2 diabetes. *Diabetes Care* 2022;**45**:965–974.
37. Oikonomou EK, Thangaraj PM, Bhatt DL, Ross JS, Young LH, Krumholz HM, et al. An explainable machine learning-based phenomapping strategy for adaptive predictive enrichment in randomized clinical trials. *NPJ Digit Med* 2023;**6**:217.
38. Banco D, Chang J, Talmor N, Wadhwa P, Mukhopadhyay A, Lu X, et al. Sex and race differences in the evaluation and treatment of young adults presenting to the emergency department with chest pain. *J Am Heart Assoc* 2022;**11**:e024199.
39. Mnatzaganian G, Hiller JE, Braitberg G, Kingsley M, Putland M, Bish M, et al. Sex disparities in the assessment and outcomes of chest pain presentations in emergency departments. *Heart* 2020;**106**:111–118.
40. Preciado SM, Sharp AL, Sun BC, Baecker A, Wu Y-L, Lee M-S, et al. Evaluating sex disparities in the emergency department management of patients with suspected acute coronary syndrome. *Ann Emerg Med* 2021;**77**:416–424.
41. Rohit M, Joseph B, Fatima R, Bruce O, Bhatt Deepak L, Albert Michelle A. Race and ethnicity considerations in patients with coronary artery disease and stroke. *J Am Coll Cardiol* 2021;**78**:2483–2492.
42. Haase R, Schlattmann P, Gueret P, Andreini D, Pontone G, Alkadhi H, et al. Diagnosis of obstructive coronary artery disease using computed tomography angiography in patients with stable chest pain depending on clinical probability and in clinically important subgroups: meta-analysis of individual patient data. *BMJ* 2019;**365**:i1945.
43. Maron DJ, Hochman JS, Reynolds HR, Bangalore S, O'Brien SM, Boden WE, et al. Initial invasive or conservative strategy for stable coronary disease. *N Engl J Med*. 2020; **382**:1395–1407.
44. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, et al. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med* 2007; **356**:1503–1516.