Journal of
Cardiovascular Development and Disease

MDPI

Article

A Fast-And-Frugal Tree for Case Finding of Diagnosed Angina Among Adults Not Receiving Cardiovascular Treatment: A UK Biobank Analysis

Abdulaziz Alrubayyi [1,2,*], Reza Fahimi [1], Habeeb Muhammad Kamal [1], Ibrahim Khalil Ja'afar [1], Mohanad Alhothify [1], Jason Madan [1] and Olalekan A. Uthman [1]

[1] Warwick Medical School, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK; reza.fahimi@warwick.ac.uk (R.F.); habeeb.muhammad-kamal@warwick.ac.uk (H.M.K.); ibrahim.ja-afar@warwick.ac.uk (I.K.J.); mohanad.alhothify@warwick.ac.uk (M.A.); j.j.madan@warwick.ac.uk (J.M.); olalekan.uthman@warwick.ac.uk (O.A.U.)
[2] College of Applied Medical Science, Jazan University, Jazan 45142, Saudi Arabia
* Correspondence: abdulaziz.alrubayyi@warwick.ac.uk

Abstract

Background: Adults with angina outside cardiovascular care may miss preventive treatment. We developed a transparent fast-and-frugal tree (FFT) for case finding in adults not receiving cardiovascular treatment. Methods: Cross-sectional analysis of UK Biobank data. The outcome was self-reported doctor-diagnosed angina. FFTs were developed from 39 candidate predictors in 294,238 participants with complete data, in all participants (primary FFT) and in those receiving neither cholesterol-lowering nor blood pressure medication (refined FFT), and compared with machine-learning models and the Pooled Cohort Equations. Results: Overall, 8970 participants (3.0%) had diagnosed angina. The primary FFT selected only these two medications. Among 63,975 untreated test participants, the refined FFT used chest pain, age, family history of heart disease and triglycerides (sensitivity 0.682; specificity 0.789; positive predictive value 0.011). Its first three cues were stable across sensitivity analyses. Logistic regression and classification and regression trees performed similarly to the refined FFT, whereas the Pooled Cohort Equations performed less well (area under the curve 0.655). Conclusions: A four-cue FFT identified two-thirds of untreated adults with diagnosed angina while flagging about one in five, and may serve as a first-stage filter for clinical assessment in primary and community care. The tool requires external and prospective validation before clinical application.

Keywords: fast-and-frugal trees; angina pectoris; case finding; UK Biobank; interpretable machine learning

Check for updates

Academic Editor: Paolo Emilio Puddu

Received: 29 July 2026
Revised: 31 October 2026
Accepted: 6 November 2026
Published: 11 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Acute chest pain is a common reason for calling Emergency Medical Services (EMS), attending the Emergency Department (ED), and seeking care at Emergency Primary Care (EPC) centres (e.g., emergency outpatient clinics, urgent treatment centres, or same-day emergency care) [1–4]. Although chest pain is the most common chief complaint of acute myocardial infarction, many non-cardiac conditions can cause similar symptoms. In the ED setting, evidence from Europe and the United States shows that chest pain accounts for one in ten ED attendances and four in ten unscheduled hospital admissions [5–7]. However, only 5.1% of ED visits for chest pain result in a diagnosis of acute coronary

syndrome (ACS), and more than half receive a diagnosis of nonspecific chest pain [8,9]. Therefore, considerable clinical and research effort has focused on rapid triage tools such as the History, Electrocardiogram (ECG), Age, Risk factors, and Troponin (HEART) score and the Prehospital HEART (PreHEART) for patients presenting with acute symptoms in the ED, EMS and EPC settings.

Chest pain presents a common but heterogeneous clinical challenge, encompassing both acute and chronic conditions that differ in presentation, risk, and care pathways. Angina is the main symptom of myocardial ischaemia and is usually caused by coronary artery disease. Stable angina carries an increased risk of acute coronary events and death, and guidelines therefore recommend that people with stable angina receive secondary prevention drugs (e.g., statins and antihypertensive treatment) to prevent events such as myocardial infarction and stroke [10]. In addition, patients with chronic stable angina often face delays in care due to limited communication between general practitioners and cardiologists and long wait times for specialist assessment, placing them at risk of adverse cardiac events [11,12]. Adults with angina who are not under cardiovascular care, whether undiagnosed or diagnosed but not treated, may therefore miss treatment that could reduce this risk. The tools in current use are not designed to identify these adults. Acute triage tools, such as HEART and PreHEART, assess patients who present with acute symptoms, whereas established cardiovascular risk scores, such as the Pooled Cohort Equations, estimate the risk of future events rather than identify people who currently have angina [13]. Identifying adults in the general population who have angina but are not yet under cardiovascular care therefore remains an unmet need that simple, transparent tools used in primary and community care could help to address.

Machine learning (ML) models have shown promising results in classifying patients with chest pain [14–16], and fast-and-frugal trees (FFTs) offer a data-driven yet interpretable alternative that relies on a small set of cues [17]. An FFT is a decision tree that asks a small number of yes-or-no questions in a fixed order and exits as soon as a decision can be made, which allows each decision to be followed and checked by the user [17]. Because FFTs need few inputs and no computing infrastructure, they can be used in settings where complex models are impractical [18,19], and transparent models of this kind can perform close to complex models in some applications [20]. Building a case-finding tool from population data, however, presents a specific difficulty. When the outcome is an existing diagnosis, the strongest predictors may be consequences of that diagnosis (e.g., treatment started after diagnosis) rather than features that help to find people who are not yet under care. A model built on such predictors is therefore likely to identify mainly people who are already known to health services. In this study, we aimed to examine how treatment-related predictors, which may follow a diagnosis, influence an FFT developed in all UK Biobank participants, and to develop a case-finding FFT for adults not receiving cardiovascular treatment, comparing its performance with standard ML models and with an established cardiovascular risk score.

2. Materials and Methods

This study was reported following the updated guidance of the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence (TRIPOD+AI) [21,22]. The TRIPOD+AI checklist is provided in Supplementary File S1.

2.1. Study Design and Data Set Description

This is a cross-sectional analysis of baseline data from the UK Biobank (UKB) cohort. The UKB is a medical database that provides comprehensive health and genetic data. The

UKB includes over 500,000 participants aged 37–73 years, recruited between 2006 and 2010. Data were collected at urban and rural assessment centres across the UK to reflect a broad range of socioeconomic and ethnic backgrounds. The dataset includes biological samples and information from questionnaires, interviews, and medical records. Further details of the study design and participant population are described elsewhere [23].

2.2. Study Population

The baseline was the date of each participant's first UKB assessment-centre visit, at which all predictors were measured. All participants with a recorded outcome and complete predictor data were eligible for the primary analysis; a history of diagnosed angina was the outcome, not an inclusion criterion.

2.3. Data Preparation and Class Imbalance

Variables were derived from the UKB baseline assessment with a scripted, reproducible pipeline. Cholesterol-lowering and blood pressure medication were derived from the sex-specific UKB medication questions (fields 6177 and 6153) combined with the medication list (field 20003); answers of "do not know" or "prefer not to answer" were treated as missing. Diabetes was defined from baseline information only. Field identifiers and all derivation rules are listed in Supplementary Tables S1 and S2.

Prior to model development, the UKB dataset underwent preprocessing steps to ensure data quality and consistency. Initially, variable types were examined and categorical variables, including sex, ethnicity, family history of heart disease, cholesterol-lowering medication, diabetes status, blood pressure medication, chest pain, type 1 and type 2 diabetes, corticosteroid use, smoking status, and physical activity, were converted to factors, with ordered levels specified for ordinal variables such as smoking and physical activity. The outcome variable, angina status, was also converted to a binary factor. We used the ROSE (Random Over-Sampling Examples) package (version 0.0-4) [24] to balance the training data, producing approximately equal numbers of angina and non-angina cases. This approach is intended to prevent the model from predicting only the majority class (non-angina) at the expense of the minority class (angina), and may therefore improve classification of the minority class [25]. The test data were not resampled, preserving the natural prevalence for evaluation.

2.4. Missing Data

Of 502,392 UKB participants, those without a recorded outcome were excluded, and the primary analysis used complete cases. Missingness for every predictor and a comparison of included and excluded participants (standardised mean differences) are reported in Supplementary Tables S5 and S6. As a sensitivity analysis, missing predictor values were imputed by multiple imputation by chained equations (five imputations, ten iterations) [26]; imputation models were fitted on the training data only, and the test-set outcome was masked during imputation.

2.5. Outcome

The outcome was a history of doctor-diagnosed angina reported at the baseline assessment, identified from the verbal interview (self-reported non-cancer illness, code 1074) or the touchscreen question on vascular and heart problems diagnosed by a doctor (code 2). The outcome is therefore prevalent at baseline. To assess its validity, self-reported angina was compared with linked hospital diagnoses (International Classification of Diseases, 10th Revision [ICD-10] code I20 for angina, and I20, I24 or I25 for ischaemic heart disease).

2.6. Variable Selection (Predictors)

Thirty-nine candidate predictors were included: demographic characteristics (age, sex and ethnicity); medical history and treatment (family history of heart disease, type 1 and type 2 diabetes, diabetes status, blood pressure medication, cholesterol-lowering medication and corticosteroid use); lifestyle (smoking status and physical activity); chest pain; cardiometabolic measures (body mass index, systolic and diastolic blood pressure, heart rate, total, high-density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol, total-to-HDL cholesterol ratio, triglycerides, random glucose and HbA1c); and blood and urine biochemistry and haematology (creatinine, urea, urinary sodium, urinary potassium and urinary creatinine, haemoglobin, haematocrit, red cell indices, red and white blood cell counts, platelet count, aspartate aminotransferase and uric acid). The candidate set was informed by a systematic review of ML methods in patients with chest pain [16]; definitions are given in Supplementary Table S2.

2.7. Sample Size

All eligible UKB participants were included; no sample size calculation was performed. The final analytic sample is reported in the participant flow (Figure 1).

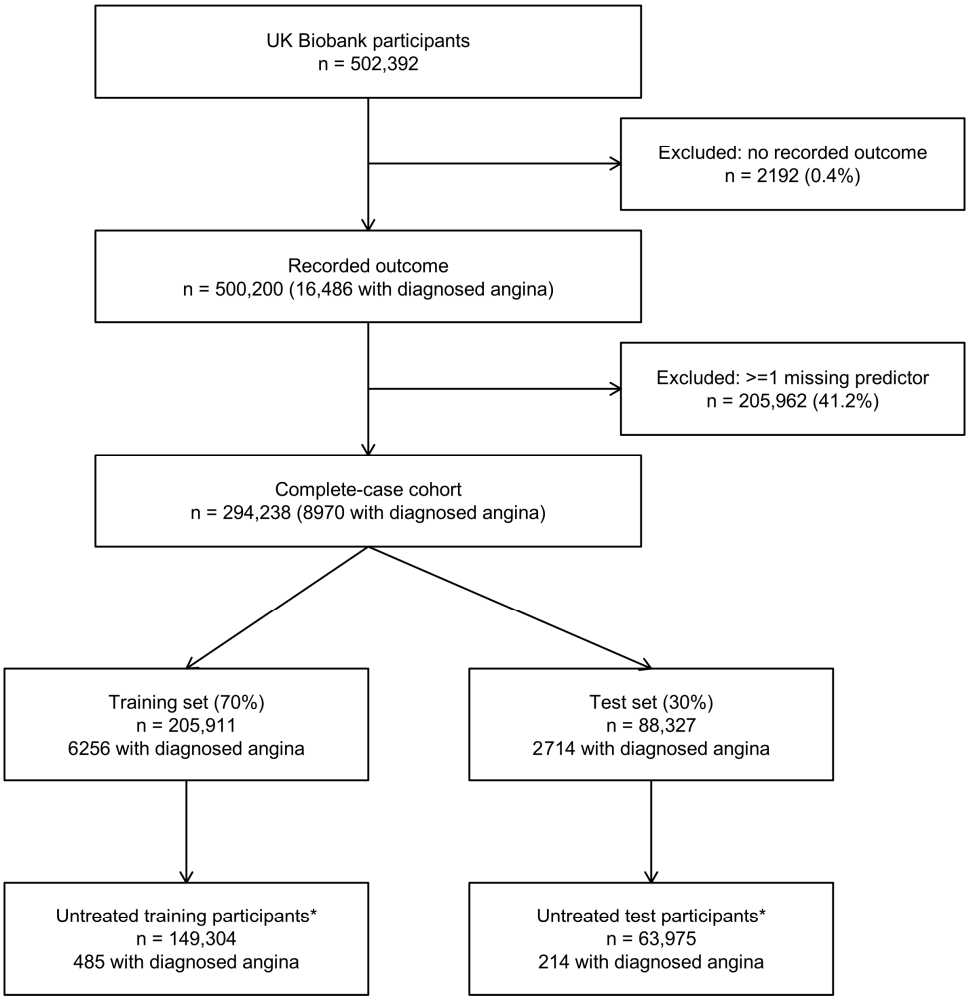


Figure 1. Participant flow, showing exclusions (with percentages of the preceding step), the complete-case cohort, the training/test split stratified by outcome and the untreated training and test participants in whom the refined FFT was developed and evaluated.

2.8. Analytical Methods

Model development was carried out in R statistical software (version 4.4.1) [27] with the FFTrees package (version 2.1.0) [17]. FFTrees implements FFTs, a supervised learning algorithm for binary classification that predicts an outcome based on a set of features (i.e., cues) for each individual [17]. Unlike traditional decision trees, FFTs use a minimal number of cues, make a binary decision at each node, and have only two branches per node. These characteristics make FFTs highly interpretable and are reported to make them relatively robust to overfitting and able to generalise well to new cases [17]. Package versions were FFTrees 2.1.0, ROSE 0.0-4, mice 3.19.0 [26], and pROC 1.19.1 [28].

The complete-case cohort was divided once into training (70%) and test (30%) sets, stratified by outcome; the same split was used for every analysis. The prevalence of diagnosed angina was 3.0% in the training set and 3.1% in the test set.

FFTs were built on the training data balanced with ROSE [24], and the tree with the highest balanced accuracy on the training data was selected; performance was evaluated once on the test data, without resampling. FFT performance was compared with logistic regression, Classification and Regression Trees (CART), random forest and support vector machines, fitted to the same ROSE-balanced training data with default settings and a 0.5 classification threshold. These models were fitted in all participants for comparison with the primary FFT and in untreated participants for comparison with the refined FFT. We report sensitivity, specificity, positive and negative predictive values, accuracy, balanced accuracy and Youden's J (sensitivity + specificity − 1) [29], with 95% confidence intervals (Wilson intervals for proportions), and the area under the receiver operating characteristic (ROC) curve (AUC) with DeLong confidence intervals for the comparison models [17,30].

To assess the influence of predictors that may follow a diagnosis of angina, FFTs were also built without the treatment cues (cholesterol-lowering and blood pressure medication) and without both the treatment cues and chest pain. Further analyses examined the effect of class balancing (no resampling) and of missing data (multiple imputation). Because the treatment cues dominated the primary tree, we developed a refined FFT for case finding in participants receiving neither cholesterol-lowering nor blood pressure medication, using the same split, candidate predictors (excluding the treatment cues) and settings. Untreated status was defined by these two medication variables only. Participants taking other cardiovascular drugs (e.g., antiplatelet agents or nitrates) were classified as untreated, because these drugs were not available as separate variables in the analysis dataset. Because undiagnosed angina cannot be observed in UK Biobank, untreated participants with diagnosed angina served as the closest available proxy for adults with angina who are outside cardiovascular care.

The refined FFT was examined with the same sensitivity analyses as the primary FFT, all in untreated participants with the same split, candidate predictors and settings: without class balancing, and with multiple imputation, in which untreated status was taken from observed answers only, so that participants with a missing medication answer were excluded regardless of the imputed value. The refined FFT was also compared with the four comparison models and the Pooled Cohort Equations in the same participants. Two further analyses addressed the nature of the outcome in this group. First, an FFT was developed without chest pain to assess the extent to which the refined FFT depends on this cue. Second, cases were classified as confirmed when angina was reported in both baseline sources (verbal interview and touchscreen) and as uncertain when it was reported in only one; the refined FFT was then evaluated after setting aside uncertain cases and redeveloped using confirmed cases only.

The Pooled Cohort Equations [13] were calculated from age, sex, ethnicity, total and HDL cholesterol, systolic blood pressure, blood pressure medication, current smoking

and diabetes for test participants aged 40 to 79 years, both in all test participants and in untreated test participants, and were evaluated by AUC and at the conventional 7.5% threshold. Because this score estimates the risk of future events, this analysis serves as a contextual benchmark rather than a like-for-like comparison.

2.9. Model Output

Each FFT produces a binary output: participants are flagged or not flagged for clinical assessment, through a sequence of cue-based decisions with an exit at every node [17]. The FFTrees algorithm constructs several candidate trees with different exit structures; all candidate trees of the primary model are reported in Supplementary Table S4.

3. Results

3.1. Participants

Of 502,392 UKB participants, 2192 (0.4%) had no recorded outcome. Of the remaining 500,200, 16,486 (3.3%) reported diagnosed angina, and 205,962 (41.2%) had at least one missing predictor, most often physical activity (19.7%), random glucose (14.4%) and HDL cholesterol (14.4%). The complete-case cohort comprised 294,238 participants, 8970 (3.0%) with diagnosed angina, divided into 205,911 training participants (6256 with angina) and 88,327 test participants (2714 with angina) (Figure 1). Included and excluded participants were similar (all absolute standardised mean differences were 0.12 or less; prevalence of diagnosed angina was 3.0% and 3.6%, respectively). Characteristics by angina status are given in Supplementary Table S3.

Self-reported angina agreed well with linked hospital records: 12,164 of 16,486 participants (73.8%) who reported diagnosed angina had a hospital diagnosis of angina (ICD-10 I20), and 14,137 (85.8%) had a diagnosis of ischaemic heart disease (I20, I24 or I25).

3.2. Primary FFT and the Influence of Treatment-Related Cues

The accuracy of each candidate cue is shown in Figure 2. The primary FFT, developed from all 39 candidate predictors, selected two cues: cholesterol-lowering medication followed by blood pressure medication (Figure 3; Table 1). On the test data it achieved a sensitivity of 0.921 (95% CI 0.910–0.931), specificity of 0.745 (0.742–0.748), positive predictive value (PPV) of 0.103 (0.099–0.107), negative predictive value (NPV) of 0.997 (0.996–0.997) and balanced accuracy of 0.833 (0.828–0.838). Without the treatment cues, the tree used chest pain, lower total and LDL cholesterol and age (sensitivity 0.816, specificity 0.781); without the treatment cues and chest pain, sensitivity fell to 0.613 (specificity 0.810). Without class balancing, the tree used cholesterol-lowering medication, and blood pressure medication combined with chest pain (sensitivity 0.894, specificity 0.835). Multiple imputation gave results close to the complete-case analysis (median sensitivity 0.914 and specificity 0.734), with the same two cues selected in all five imputations.

The trade-offs between sensitivity and specificity across all candidate trees of the primary model are reported in Supplementary Table S4.

The comparison models discriminated better than the primary FFT: AUC 0.942 (95% CI 0.937–0.946) for logistic regression, 0.937 (0.933–0.942) for random forest, 0.930 (0.925–0.935) for the support vector machine and 0.883 (0.878–0.888) for CART (Supplementary Figure S1). In test participants aged 40 to 79 years (n = 88,325), the Pooled Cohort Equations achieved an AUC of 0.746 (0.738–0.754); at the 7.5% threshold, sensitivity was 0.777 and specificity 0.599, compared with 0.921 and 0.745 for the primary FFT in the same participants. Because the PPV depends on prevalence, the primary FFT's PPV would be 0.035 at a prevalence of 1%, 0.160 at 5% and 0.286 at 10% (Supplementary Table S8).

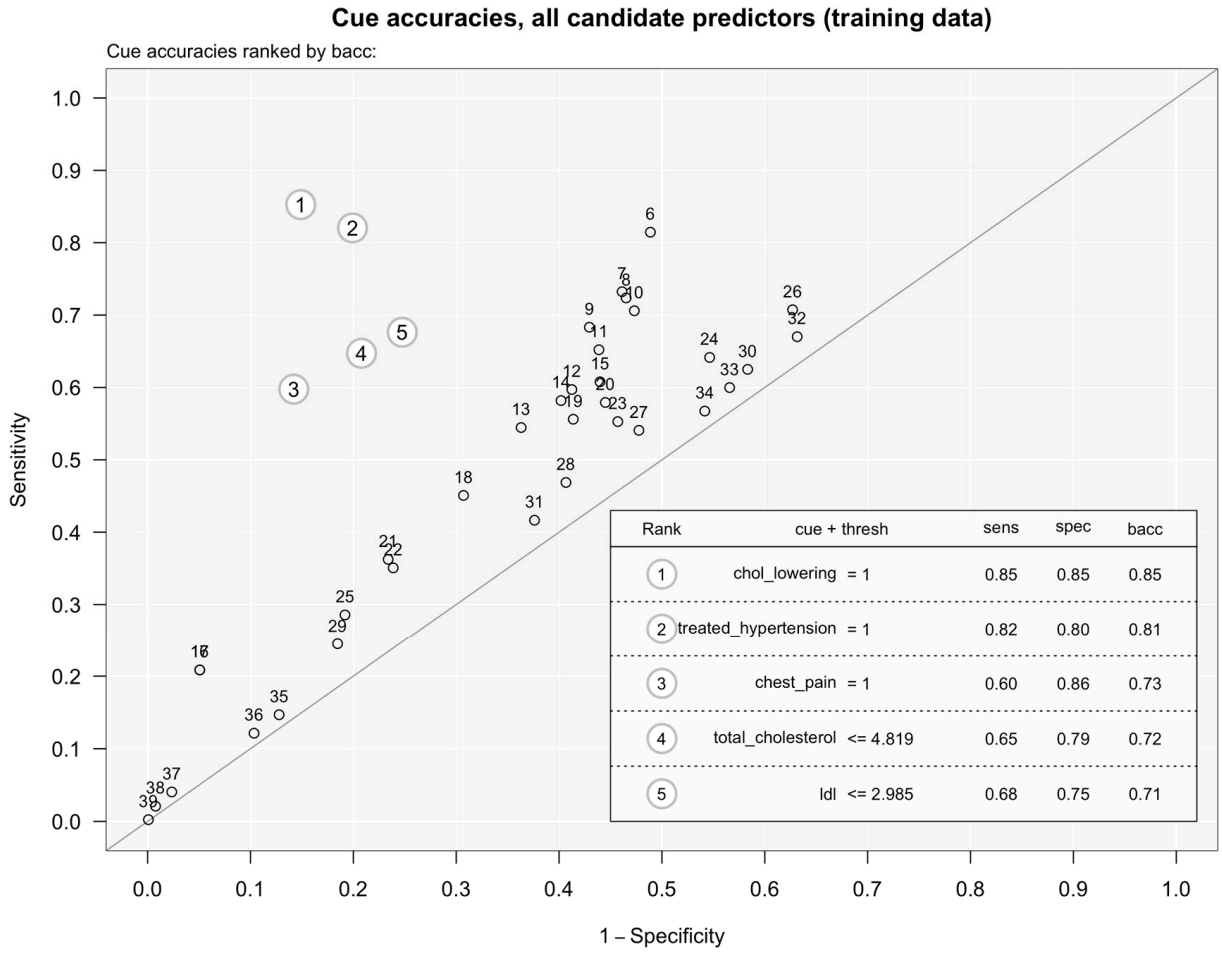


Figure 2. Accuracy of each candidate cue in the training data, shown in receiver operating characteristic space (sensitivity against 1 minus specificity) at each cue's selected threshold. Each circle is one candidate cue, numbered by its rank in balanced accuracy; the five highest-ranked cues are shown with larger circles and are listed in the inset table with their thresholds, sensitivity (sens), specificity (spec) and balanced accuracy (bacc). The diagonal line represents chance performance. chol_lowering, cholesterol-lowering medication; treated_hypertension, blood pressure medication; ldl, LDL cholesterol.

Table 1. Test-set performance of the primary FFT and sensitivity analyses (complete-case test set, $n = 88{,}327$; 2714 with diagnosed angina; multiple imputation, $n = 150{,}060$). Values are estimates with 95% confidence intervals (Wilson intervals for proportions) or, for multiple imputation, the median and range across five imputations. PPV, positive predictive value; NPV, negative predictive value; Youden's J = sensitivity + specificity − 1.

Model	Cues	Sensitivity	Specificity	PPV	NPV	Balanced Accuracy	Youden's J
All predictors (primary)	cholesterol-lowering medication; blood pressure medication	0.921 (0.910–0.931)	0.745 (0.742–0.748)	0.103 (0.099–0.107)	0.997 (0.996–0.997)	0.833 (0.828–0.838)	0.666 (0.655–0.676)
Without treatment cues	chest pain; total cholesterol; LDL cholesterol; age	0.816 (0.801–0.830)	0.781 (0.778–0.783)	0.105 (0.101–0.110)	0.993 (0.992–0.993)	0.798 (0.791–0.806)	0.596 (0.581–0.611)

Table 1. *Cont.*

Model	Cues	Sensitivity	Specificity	PPV	NPV	Balanced Accuracy	Youden's J
Without treatment cues or chest pain	total cholesterol; LDL cholesterol	0.613 (0.595–0.632)	0.810 (0.807–0.812)	0.093 (0.088–0.097)	0.985 (0.984–0.986)	0.712 (0.702–0.721)	0.423 (0.405–0.442)
All predictors, no class balancing	cholesterol-lowering medication; blood pressure medication; chest pain	0.894 (0.882–0.905)	0.835 (0.832–0.837)	0.146 (0.141–0.152)	0.996 (0.995–0.996)	0.864 (0.858–0.870)	0.728 (0.717–0.740)
All predictors, multiple imputation (median [range])	cholesterol-lowering medication; blood pressure medication	0.914 [0.914–0.915]	0.734 [0.733–0.734]	0.105 [0.105–0.105]	0.996 [0.996–0.996]	0.824 [0.824–0.824]	0.648 [0.648–0.649]

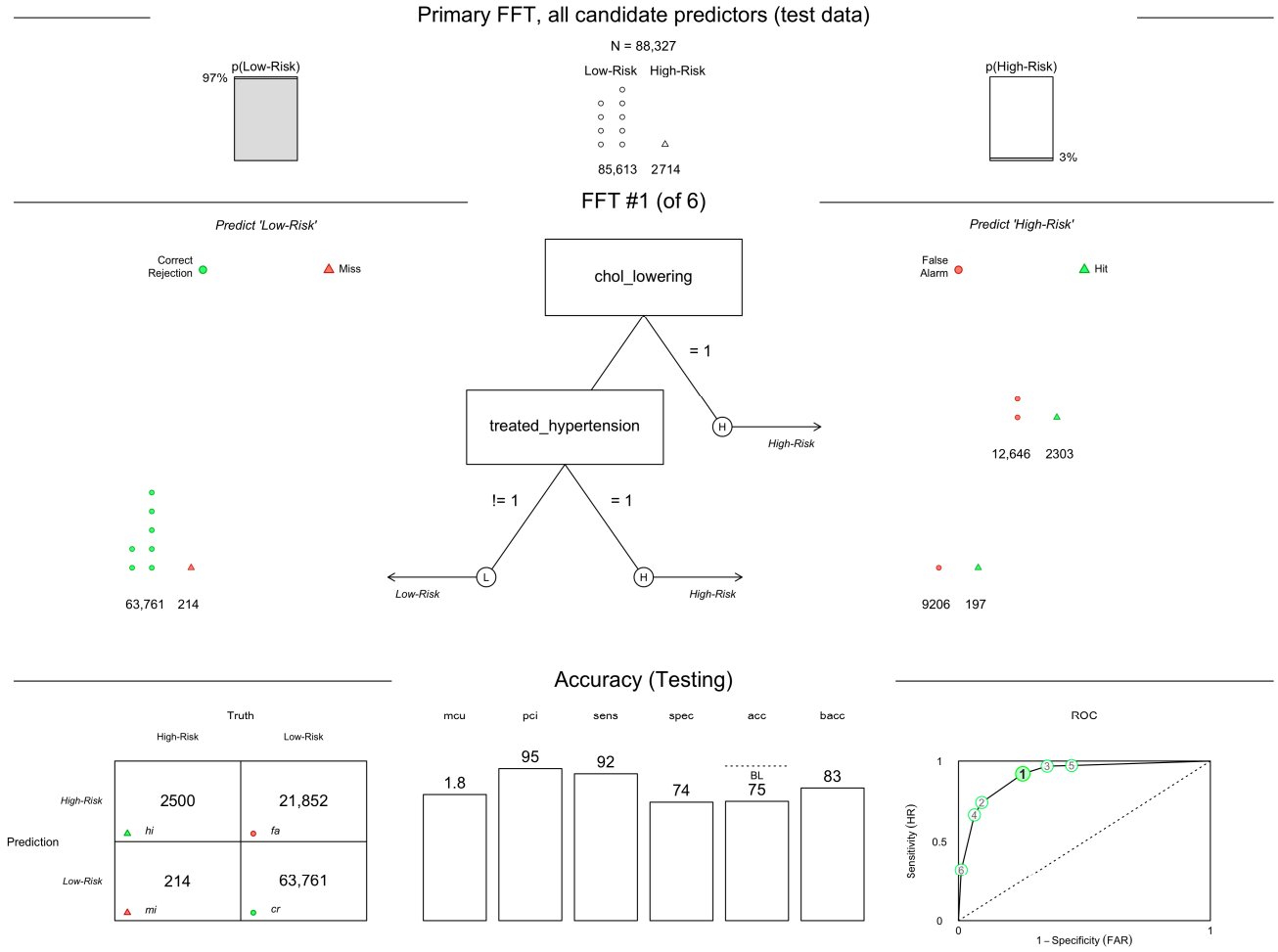


Figure 3. Primary FFT developed from all 39 candidate predictors, shown on the test data (n = 88,327; 2714 with diagnosed angina). The tree selected only cholesterol-lowering and blood pressure medication. "High-Risk" denotes participants flagged for clinical assessment. The top panel shows the number and proportion of test participants without ("Low-Risk") and with ("High-Risk") diagnosed angina. The middle panel shows the tree; at each exit (H, flagged; L, not flagged), the counts show correct rejections (green circles) and misses (red triangles) on the left, and false alarms (red circles) and hits (green triangles) on the right. The bottom panel shows the confusion table (hi, hits; mi, misses;

fa, false alarms; cr, correct rejections); the mean number of cues used (mcu); the percentage of cues ignored (pci), sensitivity (sens), specificity (spec), accuracy (acc) and balanced accuracy (bacc), in percentages, with the dashed line (BL) marking the baseline accuracy obtained by classifying all participants as "Low-Risk"; and all candidate trees in ROC space (numbered circles, with the selected tree highlighted; HR, hit rate; FAR, false alarm rate; dotted line, chance). chol_lowering, cholesterol-lowering medication; treated_hypertension, blood pressure medication.

3.3. Refined FFT for Adults Not Receiving Cardiovascular Treatment

Among the 63,975 untreated test participants, 214 (0.33%) had diagnosed angina. The refined FFT, developed in 149,304 untreated training participants (485 with angina), used four cues: participants with chest pain were flagged; those aged 57 years or younger, or without a family history of heart disease, were not flagged; and the remainder were flagged if triglycerides exceeded 1.43 mmol/L (Figure 4). It achieved a sensitivity of 0.682 (95% CI 0.617–0.741), specificity of 0.789 (0.786–0.792), PPV of 0.011 (0.009–0.013) and NPV of 0.999 (0.998–0.999), flagging 21.3% of untreated participants (Table 2). Applied to the same untreated participants, the tree developed without treatment cues in all participants had a sensitivity of 0.626 and a specificity of 0.842. The balanced accuracy of this tree (0.734, 95% CI 0.702–0.767) was similar to that of the refined FFT (0.736, 0.704–0.767).

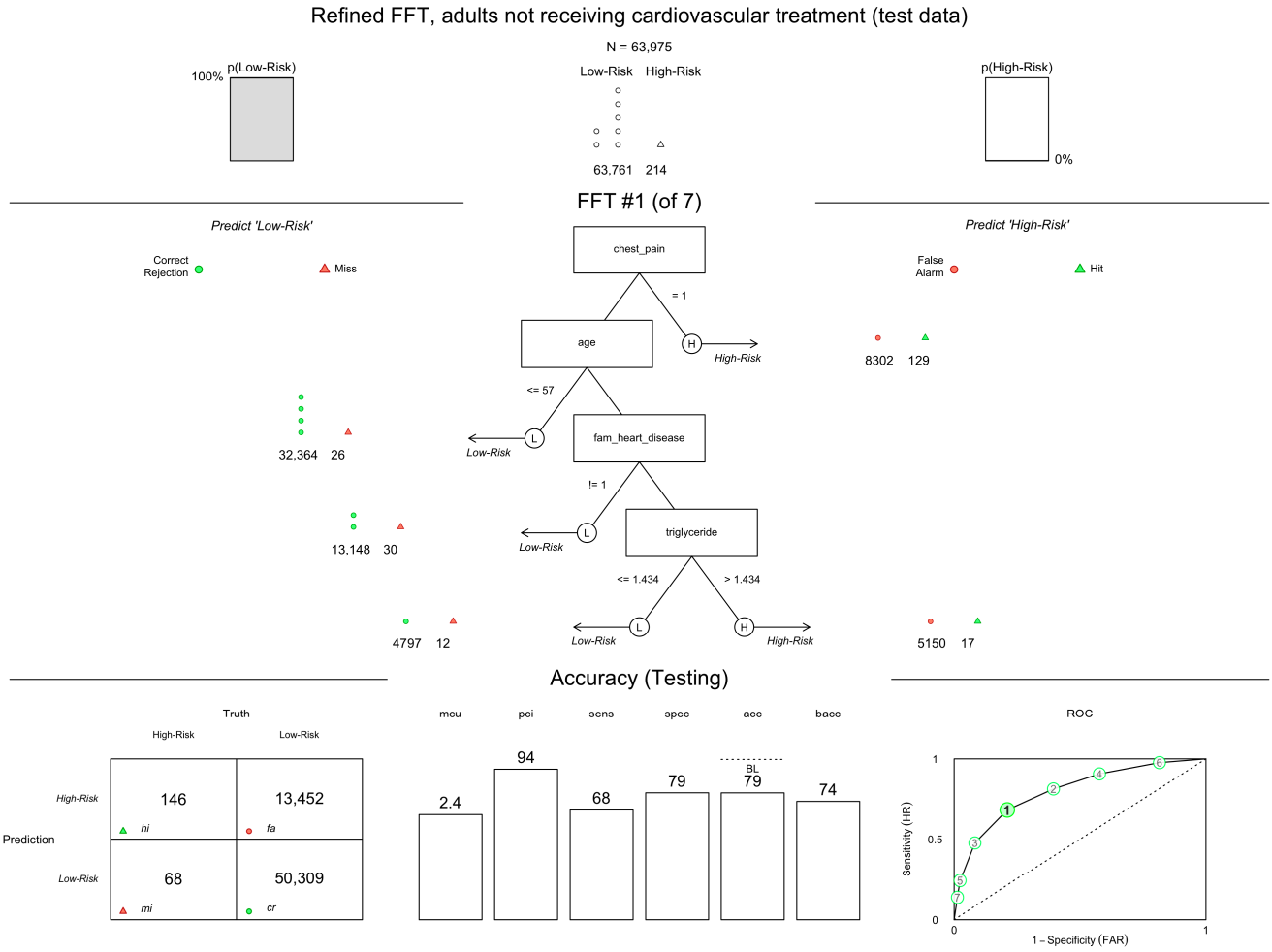


Figure 4. Refined FFT developed in adults receiving neither cholesterol-lowering nor blood pressure medication, shown on the untreated test participants (n = 63,975; 214 with diagnosed angina). "High-Risk" denotes participants flagged for clinical assessment. Panels and symbols are as described for Figure 3. fam_heart_disease, family history of heart disease; triglyceride, triglycerides (mmol/L).

Table 2. Test-set performance of the fast-and-frugal trees and comparison models in untreated participants. Panel A shows the fast-and-frugal trees and Panel B the comparison models in untreated participants (0.5 threshold; Pooled Cohort Equations at the 7.5% threshold). Multiple imputation values are the median [range] across five imputations. Uncertain cases (angina reported in only one baseline source) were set aside in the rows "uncertain cases set aside" and "redeveloped with confirmed cases only". * No participant was flagged, so the PPV could not be calculated. AUC, area under the receiver operating characteristic curve; BMI, body mass index; CART, classification and regression trees; LDL, low-density lipoprotein; NPV, negative predictive value; PPV, positive predictive value.

Model	Test Participants (Cases)	Cues	Sensitivity	Specificity	PPV	NPV	Balanced Accuracy	AUC
A. Fast-and-frugal trees								
Without treatment cues, developed in all participants	88,327 (2714)	chest pain; total cholesterol; LDL cholesterol; age	0.816 (0.801–0.830)	0.781 (0.778–0.783)	0.105 (0.101–0.110)	0.993 (0.992–0.993)	0.798 (0.791–0.806)	–
Same tree, applied to untreated participants	63,975 (214)	chest pain; total cholesterol; LDL cholesterol; age	0.626 (0.560–0.688)	0.842 (0.839–0.845)	0.013 (0.011–0.016)	0.999 (0.998–0.999)	0.734 (0.702–0.767)	–
Refined FFT, developed in untreated participants	63,975 (214)	chest pain; age; family history of heart disease; triglycerides	0.682 (0.617–0.741)	0.789 (0.786–0.792)	0.011 (0.009–0.013)	0.999 (0.998–0.999)	0.736 (0.704–0.767)	–
Refined FFT, no class balancing	63,975 (214)	chest pain; age; family history of heart disease; BMI	0.757 (0.695–0.810)	0.751 (0.747–0.754)	0.010 (0.009–0.012)	0.999 (0.999–0.999)	0.754 (0.725–0.783)	–
Refined FFT, multiple imputation (median [range])	105,524 (402)	chest pain; age; family history of heart disease; BMI	0.659 [0.657–0.659]	0.796 [0.795–0.796]	0.012 [0.012–0.012]	0.998 [0.998–0.998]	0.727 [0.726–0.727]	–
Refined FFT, uncertain cases set aside	63,945 (184)	as refined FFT	0.701 (0.631–0.763)	0.789 (0.786–0.792)	0.009 (0.008–0.011)	0.999 (0.999–0.999)	0.745 (0.712–0.778)	–
Refined FFT, redeveloped with confirmed cases only	63,945 (184)	chest pain; age; family history of heart disease; BMI	0.707 (0.637–0.768)	0.799 (0.796–0.802)	0.010 (0.008–0.012)	0.999 (0.999–0.999)	0.753 (0.720–0.786)	–
FFT without chest pain	63,975 (214)	age; family history of heart disease; triglycerides	0.696 (0.632–0.754)	0.486 (0.482–0.490)	0.005 (0.004–0.005)	0.998 (0.997–0.998)	0.591 (0.560–0.622)	–
B. Comparison models in untreated participants								
Logistic regression	63,975 (214)	–	0.752 (0.690–0.805)	0.790 (0.786–0.793)	0.012 (0.010–0.014)	0.999 (0.999–0.999)	0.771 (0.742–0.800)	0.823 (0.793–0.853)
CART	63,975 (214)	–	0.771 (0.710–0.822)	0.756 (0.753–0.760)	0.011 (0.009–0.012)	0.999 (0.999–0.999)	0.764 (0.736–0.792)	0.787 (0.756–0.817)
Random forest	63,975 (214)	–	0.000 (0.000–0.018)	1.000 (1.000–1.000)	not defined *	0.997 (0.996–0.997)	0.500 (0.500–0.500)	0.752 (0.715–0.789)

Table 2. *Cont.*

Model	Test Participants (Cases)	Cues	Sensitivity	Specificity	PPV	NPV	Balanced Accuracy	AUC
B. Comparison models in untreated participants								
Support vector machine	63,975 (214)	–	0.238 (0.186–0.300)	0.954 (0.952–0.955)	0.017 (0.013–0.022)	0.997 (0.997–0.998)	0.596 (0.567–0.624)	0.697 (0.659–0.735)
Pooled Cohort Equations, risk ≥7.5% (aged 40–79)	63,973 (214)	–	0.514 (0.447–0.580)	0.690 (0.686–0.694)	0.006 (0.005–0.007)	0.998 (0.997–0.998)	0.602 (0.569–0.636)	0.655 (0.620–0.690)

The structure of the refined FFT was largely stable in the sensitivity analyses (Table 2). Without class balancing, and in each of the five imputed datasets, the tree again began with chest pain, age and family history of heart disease, with body mass index replacing triglycerides as the fourth cue. Balanced accuracy was 0.754 (95% CI 0.725–0.783) without class balancing and 0.727 (median; range 0.726–0.727) after multiple imputation, which included 105,524 untreated test participants (402 with diagnosed angina).

In untreated participants, the comparison models discriminated less well than in all participants: the AUC was 0.823 (95% CI 0.793–0.853) for logistic regression, 0.787 (0.756–0.817) for CART, 0.752 (0.715–0.789) for random forest and 0.697 (0.659–0.735) for the support vector machine (Figure 5). At the 0.5 threshold, logistic regression and CART had balanced accuracies of 0.771 (0.742–0.800) and 0.764 (0.736–0.792), close to that of the refined FFT, whereas the random forest flagged no participants and the support vector machine had a sensitivity of 0.238. In untreated test participants aged 40 to 79 years (n = 63,973), the Pooled Cohort Equations achieved an AUC of 0.655 (0.620–0.690); at the 7.5% threshold, sensitivity was 0.514 and specificity 0.690, compared with 0.682 and 0.789 for the refined FFT in the same participants.

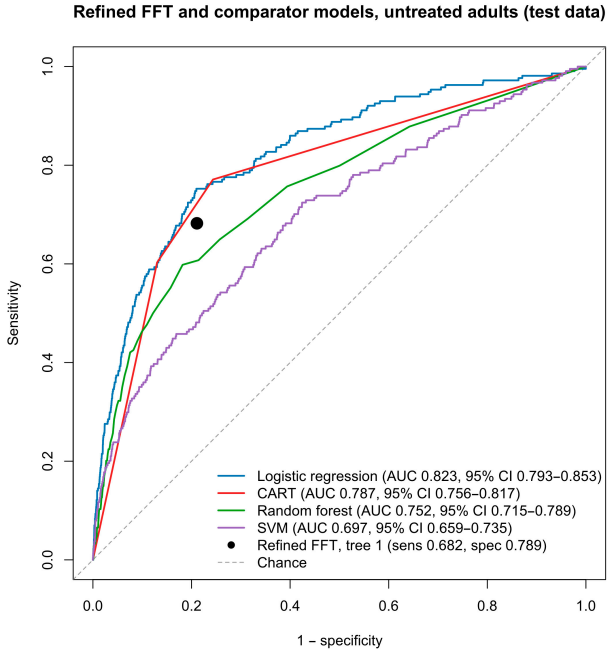


Figure 5. Receiver operating characteristic curves of the comparison models (logistic regression, CART, random forest and support vector machine) in untreated test participants (n = 63,975; 214 with diagnosed angina), with the refined FFT shown as a single operating point.

Chest pain was reported by 60.3% of untreated test participants with diagnosed angina and by 13.0% of those without. An FFT developed without chest pain used age, family history of heart disease and triglycerides; it had a sensitivity of 0.696 (95% CI 0.632–0.754) but a specificity of 0.486 (0.482–0.490), flagging 51.4% of untreated participants (balanced accuracy 0.591, 0.560–0.622; Table 2 and Supplementary Figure S2).

Untreated participants with diagnosed angina differed from treated participants with the same diagnosis (Supplementary Table S11). Of the 699 untreated cases, 36.8% had a hospital record of angina at any date, compared with 75.9% of treated cases (ischaemic heart disease, 44.6% and 89.4%), and 85.4% reported angina in both baseline sources, compared with 97.2%. Untreated cases whose angina was reported in only one source (uncertain cases) reported chest pain less often than confirmed cases (46.7% and 62.5% in the test set). When the 30 uncertain test cases were set aside, the refined FFT had a sensitivity of 0.701 (95% CI 0.631–0.763) and a specificity of 0.789 (balanced accuracy 0.745). Redeveloped with confirmed cases only, the tree again used chest pain, age (cut-off 57 years) and family history of heart disease, with body mass index as the fourth cue (sensitivity 0.707, specificity 0.799, balanced accuracy 0.753) (Table 2).

4. Discussion

In adults not receiving cardiovascular treatment, the population in which case finding is likely to be most useful, a refined four-cue FFT (chest pain, age, family history of heart disease and triglycerides) identified two-thirds of those with diagnosed angina while leaving about four in five participants unflagged. Its PPV was low, which largely reflects the low prevalence of diagnosed angina in this group. This refined tree was needed because, in UK Biobank, an FFT trained to classify diagnosed angina in all participants selected treatment variables: the primary tree identified participants receiving cholesterol-lowering or blood pressure medication, who are largely those already diagnosed. Standard ML models discriminated better than the primary FFT in all participants, but in untreated participants logistic regression and CART performed similarly to the refined FFT, and random forest and the support vector machine performed less well at the 0.5 threshold. Therefore, the contribution of the refined FFT lies in its transparency and ease of use rather than in higher accuracy. The first three cues of the refined FFT (chest pain, age and family history of heart disease) were the same in every sensitivity analysis, whereas the fourth, metabolic cue alternated between triglycerides and body mass index.

The dominance of treatment cues illustrates a general issue in models of prevalent disease: variables that follow a diagnosis, such as treatment, treatment-modified biomarkers and symptoms that prompted diagnosis, can classify the outcome well while carrying little information for finding undiagnosed cases. Excluding the treatment cues did not remove this effect entirely, because the resulting tree used lower total and LDL cholesterol, consistent with statin treatment after diagnosis. Developing the case-finding tree in untreated adults avoided these cues and produced a clinically interpretable rule. However, the advantage of this approach lies in interpretability rather than performance. In the same untreated participants, the tree developed without treatment cues had almost identical balanced accuracy (0.734 vs. 0.736), but it relied on the treatment-modified cholesterol cues described above, whereas the refined FFT avoids these measures. This illustrates, in practice, the principle that models for finding undiagnosed disease should not rely on information that follows a diagnosis.

Chest pain was the first cue of the refined FFT, which raises a specific form of circularity. In adults with diagnosed angina, chest pain is the symptom through which the condition usually presents, and participants who know their diagnosis may report it more readily. In the untreated test set, 60.3% of participants with diagnosed angina reported chest pain,

compared with 13.0% of those without. The performance of the refined FFT may therefore overestimate its ability to identify undiagnosed angina. However, chest pain did not define the outcome in this population. Of the 45,963 complete-case participants who reported chest pain, only 5380 (11.7%) had diagnosed angina, and among untreated test participants flagged at the first cue, only 129 of 8431 (1.5%) had diagnosed angina. This is consistent with clinical evidence that chest pain is common and frequently non-cardiac, with causes that include respiratory, gastrointestinal, musculoskeletal and psychological conditions (e.g., anxiety and panic disorder) [8,9]. Without chest pain, an FFT developed in the same untreated adults kept a similar sensitivity (0.696) but flagged about half of the participants (specificity 0.486), suggesting that chest pain provides most of the ability of the tree to separate participants with and without diagnosed angina. Cases whose angina was reported in only one baseline source reported chest pain less often than confirmed cases (46.7% and 62.5%), suggesting that misreported diagnoses are unlikely to explain the strength of this cue. Chest pain therefore acts as a common but non-specific first step, and participants flagged by the tree would require clinical assessment to establish the cause of their symptoms.

Previous studies in this field provide limited direct context for interpreting our findings, since existing clinical risk scores such as HEART, PreHEART, and the Troponin-only Manchester Acute Coronary Syndromes (T-MACS) score were developed and validated for a different task: acute rule-out of ACS in patients presenting with active chest pain, rather than identification of diagnosed angina in a general population [31–34]. Reported performance for these tools in prehospital and emergency primary care settings has been mixed and setting-dependent, with sensitivity and negative predictive value varying according to the population studied and the outcome definition used [32–38]. These scores require ECG and troponin measurements, which carry strong discriminatory value [35] but are not available in UK Biobank. As a contextual benchmark, we used the Pooled Cohort Equations, which achieved an AUC of 0.746 in all participants and 0.655 in untreated participants and, at the 7.5% threshold, had lower sensitivity and specificity than the treatment-based primary FFT and the refined FFT, respectively. However, because this score predicts future events rather than prevalent angina, these comparisons should not be interpreted as evidence of superiority. A more directly comparable study is that of Mousavi et al., who applied multiple ML algorithms to predict angina in a community-dwelling elderly population, reporting sensitivities ranging from 65.5% to 79.2% and specificities ranging from 62.1% to 77.3% [39]. The refined FFT reached a comparable sensitivity (0.682) and specificity (0.789) with four routinely available cues, and its reasoning can be inspected directly at every step, which may support trust and usability in case finding. Its cues comprise history items and a routinely measured blood lipid, which can be collected without specialist equipment.

We focused on diagnosed angina because of its relevance to case finding. As noted in the Introduction, stable angina carries an increased risk of acute coronary events and death, and its management aims both to control symptoms and to improve long-term morbidity and mortality [10]. Angina is also heterogeneous, as chest pain may reflect obstructive coronary disease, non-obstructive atherosclerosis, coronary microvascular dysfunction or non-cardiac mechanisms, which have different diagnostic and prognostic implications [40]. A self-reported, doctor-diagnosed outcome cannot distinguish between these syndromes, or between stable and unstable angina, so the present models address diagnosed angina as a whole.

Recent advances in the early diagnosis of myocardial infarction (MI), particularly the use of high-sensitivity cardiac troponin (hs-cTn) assays, have altered the clinical presentation and definition of unstable angina (UA) [41,42]. In addition, the transition point from unstable to stable angina is poorly defined [43]. Clinically, UA is typically identified by features such as prolonged angina lasting more than 20 min at rest, new onset of severe

angina, increasing frequency or duration of angina episodes, or angina occurring shortly after a recent MI [44]. Individuals with diagnosed angina may fall across a broad spectrum of ischaemic risk, from stable disease to features suggestive of UA or an alternative diagnosis [45–49]. Individuals flagged by a case-finding tool would therefore require clinical assessment to establish the nature of their symptoms.

Interpretation of predictive values requires care. The high NPV of all models largely reflects the low prevalence of diagnosed angina, whereas PPVs were modest: 0.103 for the primary FFT and 0.011 for the refined FFT in untreated adults, where about 1 in 93 flagged participants had diagnosed angina and 21.3% of participants were flagged. Without the tree, about 1 in 300 untreated participants had diagnosed angina (prevalence 0.33%). Assessing flagged participants would therefore find about three times as many cases per assessment as assessing participants at random, and 78.7% of untreated participants would not require further assessment. The cost of this gain is that about one in three participants with diagnosed angina was not flagged. The refined FFT is therefore best viewed as a transparent first-stage filter that prioritises adults for clinical assessment, not as a diagnostic test, and its value in practice will depend on the capacity for, and cost of, follow-up assessment. The trees were selected by balanced accuracy, which gives equal weight to sensitivity and specificity. Where missed cases are of greater concern than additional assessments, a tree with higher sensitivity and lower specificity could be preferred; the trade-offs of all candidate trees of the primary model are reported in Supplementary Table S4.

The refined FFT may be most useful where adults outside regular cardiovascular care are already in contact with health services, such as in primary care. Community paramedicine (also called the primary care paramedic model) is one possible future setting. In this model, which is practised in many countries (e.g., Australia, Canada and the UK), trained paramedics assess patients in the community and can refer them to primary care, or manage them without hospital transport, when appropriate [50–55]. Because the refined FFT uses few and simple cues, it could in principle be applied during such contacts, although measuring triglycerides requires a blood test. This setting was not studied here, and UK Biobank participants were assessed in a research setting rather than during paramedic encounters, so any use in this setting would require dedicated validation. External and prospective validation, ideally in primary-care populations, is required before any clinical use, together with an assessment of the costs and benefits of the additional clinical assessments that a case-finding tool would generate.

This study has several strengths. The analysis was based on a large, well-characterised cohort and used a transparent modelling approach, and the self-reported outcome was validated against linked hospital records. In addition, variable derivation and analysis followed a reproducible pipeline, with sensitivity analyses for treatment-related cues, class balancing and missing data. Furthermore, the FFTs were benchmarked against standard ML models and the Pooled Cohort Equations, both in all participants and in the untreated population.

This study also has several limitations. The outcome was prevalent self-reported diagnosed angina, which means that the models classify existing diagnoses rather than predict future disease, and recall error cannot be excluded, although 73.8% of self-reported cases were confirmed in hospital records. In addition, untreated participants with diagnosed angina, a group less often confirmed by hospital records than treated cases, served as a proxy for adults with angina outside cardiovascular care; therefore, whether the refined FFT can identify undiagnosed angina requires prospective evaluation. Furthermore, validation was internal, and UKB participants are healthier than the general population [56], so external and prospective validation is needed, as performance in practice is often lower [57]. Moreover, 41.2% of participants with a recorded outcome were excluded because of missing predictors, although included and excluded participants were similar and multiple impu-

tation gave consistent results for both the primary and the refined FFTs. The low PPV in untreated adults also implies that a large proportion of flagged participants would be false positives. Another limitation is that ECG and troponin were not available as predictors in the UKB, and although their absence may enhance feasibility in settings with limited access to diagnostic testing [58], it limits comparability with established risk scores that incorporate biomarkers and ECG findings. Finally, residual effects of treatment and of diagnosis on predictor values cannot be excluded, including more frequent reporting of chest pain by participants aware of their diagnosis.

5. Conclusions

In adults not receiving cardiovascular treatment, a transparent four-cue fast-and-frugal tree (chest pain, age, family history of heart disease and triglycerides) identified two-thirds of those with diagnosed angina while flagging about one in five participants. Because its positive predictive value was low, the tree may serve as a first-stage filter that prioritises adults for clinical assessment in primary and community care, rather than as a diagnostic test. Its first three cues were stable across sensitivity analyses, and in untreated adults it performed similarly to logistic regression and CART while remaining easy to interpret. However, this tool requires external and prospective validation before clinical application.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/jcdd13110587/s1. File S1: TRIPOD+AI checklist. Table S1: UK Biobank field identifiers for the outcome and the 39 candidate predictors. Table S2: Derivation and recoding of each variable. Table S3: Characteristics of the complete-case cohort by angina status (n = 294,238). Table S4: Test-set performance of all candidate trees of the primary FFT (n = 88,327). Table S5: Missing values for each candidate predictor among participants with a recorded outcome (n = 500,200). Table S6: Participants included in (complete cases) and excluded from (any missing predictor) the primary analysis. Table S7: Multiple imputation sensitivity analysis: FFT selected in each imputed dataset (test set, n = 150,060). Table S8: Predictive values of the primary FFT at different prevalences of diagnosed angina. Table S9: Pooled Cohort Equations benchmark (test participants aged 40 to 79 years). Table S10: Derivation of the analysis dataset: corrections relative to the originally submitted dataset. Table S11: Characteristics of participants with diagnosed angina, by treatment status (complete-case cohort). Figure S1: Receiver operating characteristic curves of the comparison models (logistic regression, CART, random forest and support vector machine) in all test participants (n = 88,327; 2714 with diagnosed angina), with the primary FFT shown as a single operating point. Figure S2: FFT developed without chest pain in adults receiving neither cholesterol-lowering nor blood pressure medication, shown on the untreated test participants (n = 63,975; 214 with diagnosed angina). Ref. [59] is included in the Supplementary Materials.

Author Contributions: Conceptualization, A.A. and O.A.U.; methodology, A.A. and O.A.U.; validation, R.F., J.M. and O.A.U.; formal analysis, A.A. and O.A.U.; investigation, A.A.; data curation, A.A. and O.A.U.; writing—original draft preparation, A.A.; writing—review and editing, R.F., H.M.K., I.K.J., M.A., J.M. and O.A.U.; visualization, A.A. and O.A.U.; supervision, O.A.U. and J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Biomedical and Scientific Research Ethics Committee (BSREC) of the University of Warwick (6 May 2021, reference number BSREC 104/20-21).

Informed Consent Statement: UK Biobank participants provided written informed consent at recruitment. All data were de-identified before release to the research team. No patients or public representatives were involved in any aspect of this study.

Data Availability Statement: Restrictions apply to the availability of these data. Data were obtained from UK Biobank and are available to researchers with the permission of UK Biobank, subject to

a data access application (https://www.ukbiobank.ac.uk); this research used application number 71756. The analysis code is available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACS	acute coronary syndrome
AUC	area under the receiver operating characteristic curve
BMI	body mass index
CART	classification and regression trees
CI	confidence interval
ECG	Electrocardiogram
ED	emergency department
EMS	emergency medical services
EPC	emergency primary care
FFT	fast-and-frugal tree
HbA1c	glycated haemoglobin
HDL	high-density lipoprotein
HEART	History, ECG, Age, Risk factors and Troponin score
hs-cTn	high-sensitivity cardiac troponin
ICD-10	International Classification of Diseases, 10th revision
LDL	low-density lipoprotein
MI	myocardial infarction
ML	machine learning
NPV	negative predictive value
PPV	positive predictive value
PreHEART	Prehospital HEART score
ROC	receiver operating characteristic
ROSE	Random Over-Sampling Examples
T-MACS	Troponin-only Manchester Acute Coronary Syndromes decision aid
TRIPOD+AI	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis, updated for artificial intelligence
UA	unstable angina
UKB	UK Biobank

References

1. Ibsen, S.; Lindskou, T.A.; Nickel, C.H.; Kløjgård, T.; Christensen, E.F.; Søvsø, M.B. Which symptoms pose the highest risk in patients calling for an ambulance? A population-based cohort study from Denmark. *Scand. J. Trauma. Resusc. Emerg. Med.* **2021**, *29*, 59. [CrossRef] [PubMed]
2. Alghamdi, A.; Howard, L.; Reynard, C.; Moss, P.; Jarman, H.; Mackway-Jones, K.; Carley, S.; Body, R. Enhanced triage for patients with suspected cardiac chest pain: The History and Electrocardiogram-only Manchester Acute Coronary Syndromes decision aid. *Eur. J. Emerg. Med.* **2019**, *26*, 356–361. [CrossRef] [PubMed]
3. Blinkenberg, J.; Pahlavanyali, S.; Hetlevik, Ø.; Sandvik, H.; Hunskaar, S. General practitioners' and out-of-hours doctors' role as gatekeeper in emergency admissions to somatic hospitals in Norway: Registry-based observational study. *BMC Health Serv. Res.* **2019**, *19*, 568. [CrossRef] [PubMed]
4. Burman, R.A.; Zakariassen, E.; Hunskaar, S. Management of chest pain: A prospective study from Norwegian out-of-hours primary care. *BMC Fam. Pract.* **2014**, *15*, 51. [CrossRef] [PubMed]
5. Goodacre, S.; Cross, E.; Arnold, J.; Angelini, K.; Capewell, S.; Nicholl, J. The health care burden of acute chest pain. *Heart* **2005**, *91*, 229–230. [CrossRef] [PubMed]
6. Goodacre, S.; Thokala, P.; Carroll, C.; Stevens, J.W.; Leaviss, J.; Al Khalaf, M.; Collinson, P.; Morris, F.; Evans, P.; Wang, J. Systematic review, meta-analysis and economic modelling of diagnostic strategies for suspected acute coronary syndrome. *Health Technol. Assess.* **2013**, *17*, 1–188. [CrossRef] [PubMed]
7. Virani, S.S.; Alonso, A.; Benjamin, E.J.; Bittencourt, M.S.; Callaway, C.W.; Carson, A.P.; Chamberlain, A.M.; Chang, A.R.; Cheng, S.; Delling, F.N. Heart disease and stroke statistics—2020 update: A report from the American Heart Association. *Circulation* **2020**, *141*, e139–e596. [CrossRef] [PubMed]

8. Gulati, M.; Levy, P.D.; Mukherjee, D.; Amsterdam, E.; Bhatt, D.L.; Birtcher, K.K.; Blankstein, R.; Boyd, J.; Bullock-Palmer, R.P.; Conejo, T.; et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* **2021**, *144*, e368–e454. [CrossRef] [PubMed]
9. Hsia, R.Y.; Hale, Z.; Tabas, J.A. A National Study of the Prevalence of Life-Threatening Diagnoses in Patients With Chest Pain. *JAMA Intern. Med.* **2016**, *176*, 1029–1032. [CrossRef] [PubMed]
10. National Institute For Health and Care Excellence (NICE). *Stable Angina: Management*; Clinical guideline; National Institute For Health and Care Excellence (NICE): London, UK, 2016.
11. van den Bulk, S.; Spoelman, W.A.; van Dijkman, P.R.; Numans, M.E.; Bonten, T.N.; Leiden University Medical Center. Non-acute chest pain in primary care; referral rates, communication and guideline adherence: A cohort study using routinely collected health data. *BMC Prim. Care* **2022**, *23*, 336. [CrossRef] [PubMed]
12. Afgani, F.J.; Brenna, C.T.; Hanneman, K.; Holzapfel, S.; Burton, C.; Harvey, P.J.; Nguyen, E.T. Cardiac Link: A retrospective cohort study evaluating a clinical pathway for expedited cardiology referral. *CMAJ Open* **2022**, *10*, E1000–E1007. [CrossRef] [PubMed]
13. Goff, D.C.; Lloyd-Jones, D.M.; Bennett, G.; Coady, S.; D'Agostino, R.B.; Gibbons, R.; Greenland, P.; Lackland, D.T.; Levy, D.; O'Donnell, C.J.; et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. *Circulation* **2014**, *129*, S49–S73. [CrossRef] [PubMed]
14. Wibring, K.; Lingman, M.; Herlitz, J.; Ashfaq, A.; Bång, A. Development of a prehospital prediction model for risk stratification of patients with chest pain. *Am. J. Emerg. Med.* **2022**, *51*, 26–31. [CrossRef] [PubMed]
15. Alrubayyi, A.; Fahimi, R.; Hassan, M.; Madan, J.; Uthman, O.A. Current Clinical Practices and Future Perspectives for Primary Healthcare Use of Point-of-Care Devices: A Scoping Review. *Int. J. Clin. Pract.* **2025**, *2025*, 4742851. [CrossRef]
16. Stewart, J.; Lu, J.; Goudie, A.; Bennamoun, M.; Sprivulis, P.; Sanfillipo, F.; Dwivedi, G. Applications of machine learning to undifferentiated chest pain in the emergency department: A systematic review. *PLoS ONE* **2021**, *16*, e0252612. [CrossRef] [PubMed]
17. Phillips, N.D.; Neth, H.; Woike, J.K.; Gaissmaier, W. FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. *Judgm. Decis. Mak.* **2017**, *12*, 344–368. [CrossRef]
18. Gigerenzer, G.; Gaissmaier, W. Heuristic Decision Making. *Annu. Rev. Psychol.* **2011**, *62*, 451–482. [CrossRef] [PubMed]
19. Marewski, J.N.; Gigerenzer, G. Heuristic decision making in medicine. *Dialogues Clin. Neurosci.* **2012**, *14*, 77–89. [CrossRef] [PubMed]
20. Katsikopoulos, K.V.; Şimşek, Ö.; Buckmann, M.; Gigerenzer, G. *Classification in the Wild: The Science and Art of Transparent Decision Making*; The MIT Press: Cambridge, MA, USA, 2021; ISBN 978-0-262-36322-8.
21. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Ghassemi, M.; Liu, X.; Reitsma, J.B.; van Smeden, M.; et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **2024**, *385*, e078378. [CrossRef] [PubMed]
22. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ* **2015**, *350*, g7594. [CrossRef] [PubMed]
23. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; 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. [CrossRef] [PubMed]
24. Lunardon, N.; Menardi, G.; Torelli, N. ROSE: A package for binary imbalanced learning. *R J.* **2014**, *6*, 79–89. [CrossRef]
25. Blagus, R.; Lusa, L. Joint use of over-and under-sampling techniques and cross-validation for the development and assessment of prediction models. *BMC Bioinform.* **2015**, *16*, 363. [CrossRef] [PubMed]
26. van Buuren, S.; Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Soft.* **2011**, *45*, 1–67. [CrossRef]
27. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2024.
28. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M.; Siegert, S.; Doering, M.; Billings, Z. pROC: Display and Analyze ROC Curves. 2026. Available online: https://cran.r-project.org/web/packages/pROC/index.html (accessed on 15 October 2026).
29. Kallner, A. Formulas. In *Laboratory Statistics*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 1–140, ISBN 978-0-12-814348-3.
30. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845. [CrossRef] [PubMed]
31. Six, A.J.; Backus, B.E.; Kelder, J.C. Chest pain in the emergency room: Value of the HEART score. *Neth. Heart J.* **2008**, *16*, 191–196. [CrossRef] [PubMed]

32. Alghamdi, A.; Hann, M.; Carlton, E.; Cooper, J.G.; Cook, E.; Foulkes, A.; Siriwardena, A.N.; Phillips, J.; Thompson, A.; Bell, S.; et al. Diagnostic Accuracy of Clinical Pathways for Suspected Acute Myocardial Infarction in the Out-of-Hospital Environment. *Ann. Emerg. Med.* **2023**, *82*, 439–448. [CrossRef] [PubMed]
33. Cooper, J.G.; Donaldson, L.A.; Coutts, A.J.; Body, R.; Mills, N.L. Prehospital T-MACS and HEART scores in the prediction of myocardial infarction: A prospective evaluation. *Emerg. Med. J.* **2024**, *41*, 255–256. [CrossRef] [PubMed]
34. Cooper, J.G.; Ferguson, J.; Donaldson, L.A.; Black, K.M.; Livock, K.J.; Horrill, J.L.; Davidson, E.M.; Scott, N.W.; Lee, A.J.; Fujisawa, T. Performance of a prehospital HEART score in patients with possible myocardial infarction: A prospective evaluation. *Emerg. Med. J.* **2023**, *40*, 474–481. [CrossRef] [PubMed]
35. Sagel, D.; Vlaar, P.J.; van Roosmalen, R.; Waardenburg, I.; Nieuwland, W.; Lettinga, R.; van Barneveld, R.; Jorna, E.; Kijlstra, R.; van Well, C.; et al. Prehospital risk stratification in patients with chest pain. *Emerg. Med. J.* **2021**, *38*, 814–819. [CrossRef] [PubMed]
36. Johannessen, T.R.; Melessen, I.M.B.; Vallersnes, O.M.; Manten, A.; Halvorsen, S.; Atar, D.; Harskamp, R.E. Adequately identifying low-risk chest pain in emergency primary care: Evaluating the performance of preHEAR(T) based on two European cohorts. *Open Heart* **2025**, *12*, e003362. [CrossRef] [PubMed]
37. Demandt, J.P.A.; Koks, A.; Sagel, D.; Haest, R.; Heijmen, E.; Thijssen, E.; el Farissi, M.; Eerdekens, R.; van der Harst, P.; van 't Veer, M.; et al. External validation of the preHEART score and comparison with current clinical risk scores for prehospital risk assessment in patients with suspected NSTE-ACS. *Emerg. Med. J.* **2024**, *41*, 610–616. [CrossRef] [PubMed]
38. Johannessen, T.R.; Atar, D.; Vallersnes, O.M.; Larstorp, A.C.K.; Mdala, I.; Halvorsen, S. Comparison of a single high-sensitivity cardiac troponin T measurement with the HEART score for rapid rule-out of acute myocardial infarction in a primary care emergency setting: A cohort study. *BMJ Open* **2021**, *11*, e046024. [CrossRef] [PubMed]
39. Mousavi, S.; Jalalian, Z.; Afrashteh, S.; Farhadi, A.; Nabipour, I.; Larijani, B. Machine learning-based predictive modeling of angina pectoris in an elderly community-dwelling population: Results from the PoCOsteo study. *PLoS ONE* **2025**, *20*, e0329023. [CrossRef] [PubMed]
40. Pizzi, C.; Bugiardini, R. Further insights into syndrome X. *Heart* **2010**, *96*, 1865–1867. [CrossRef] [PubMed]
41. D'Souza, M.; Sarkisian, L.; Saaby, L.; Poulsen, T.S.; Gerke, O.; Larsen, T.B.; Diederichsen, A.C.; Jangaard, N.; Diederichsen, S.Z.; Hosbond, S. Diagnosis of unstable angina pectoris has declined markedly with the advent of more sensitive troponin assays. *Am. J. Med.* **2015**, *128*, 852–860. [CrossRef] [PubMed]
42. Braunwald, E.; Morrow, D.A. Unstable angina: Is it time for a requiem? *Circulation* **2013**, *127*, 2452–2457. [PubMed]
43. Rao, S.V.; O'Donoghue, M.L.; Ruel, M.; Rab, T.; Tamis-Holland, J.E.; Alexander, J.H.; Baber, U.; Baker, H.; Cohen, M.G.; Cruz-Ruiz, M. 2025 ACC/AHA/ACEP/NAEMSP/SCAI guideline for the management of patients with acute coronary syndromes: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* **2025**, *151*, e771–e862. [CrossRef] [PubMed]
44. Byrne, R.A.; Rossello, X.; Coughlan, J.J.; Barbato, E.; Berry, C.; Chieffo, A.; Claeys, M.J.; Dan, G.-A.; Dweck, M.R.; Galbraith, M. 2023 ESC guidelines for the management of acute coronary syndromes. *Eur. Heart J.* **2023**, *44*, 3720–3826. [CrossRef] [PubMed]
45. Boothroyd, L.J.; Lambert, L.J.; Segal, E.; Ross, D.; Kouz, S.; Maire, S.; Harvey, R.; Xiao, Y.; Brown, K.A.; Nasmith, J.; et al. Comparison of Outcomes of Ambulance Users and Nonusers in ST Elevation Myocardial Infarction. *Am. J. Cardiol.* **2014**, *114*, 1289–1294. [CrossRef] [PubMed]
46. Lavery, T.; Greenslade, J.H.; Parsonage, W.A.; Hawkins, T.; Dalton, E.; Hammett, C.; Cullen, L. Factors influencing choice of pre-hospital transportation of patients with potential acute coronary syndrome: An observational study. *Emerg. Med. Australas.* **2017**, *29*, 210–216. [CrossRef] [PubMed]
47. Thang, N.D.; Karlson, B.W.; Bergman, B.O.; Santos, M.; Karlsson, T.; Bengtson, A.; Johanson, P.; Rawshani, A.; Herlitz, J. Characteristics of and outcome for patients with chest pain in relation to transport by the emergency medical services in a 20-year perspective. *Am. J. Emerg. Med.* **2012**, *30*, 1788–1795. [CrossRef] [PubMed]
48. Thuresson, M.; Jarlöv, M.B.; Lindahl, B.; Svensson, L.; Zedigh, C.; Herlitz, J. Factors that influence the use of ambulance in acute coronary syndrome. *Am. Heart J.* **2008**, *156*, 170–176. [CrossRef] [PubMed]
49. Thygesen, K.; Alpert, J.S.; Jaffe, A.S.; Chaitman, B.R.; Bax, J.J.; Morrow, D.A.; White, H.D.; ESC Scientific Document Group. Fourth universal definition of myocardial infarction (2018). *Eur. Heart J.* **2018**, *40*, 237–269. [CrossRef] [PubMed]
50. Drennan, I.R.; Blanchard, I.E.; Buick, J.E. Opportunity for change: Is it time to redefine the role of paramedics in healthcare? *Can. J. Emerg. Med.* **2021**, *23*, 139–140. [CrossRef] [PubMed]
51. Oosterwold, J.; Sagel, D.; Berben, S.; Roodbol, P.; Broekhuis, M. Factors influencing the decision to convey or not to convey elderly people to the emergency department after emergency ambulance attendance: A systematic mixed studies review. *BMJ Open* **2018**, *8*, e021732. [CrossRef] [PubMed]
52. Shannon, B.; Eaton, G.; Lanos, C.; Leyenaar, M.; Nolan, M.; Bowles, K.; Williams, B.; O'Meara, P.; Wingrove, G.; Heffern, J.; et al. The development of community paramedicine; a restricted review. *Health Soc. Care Community* **2022**, *30*, e3547–e3561. [CrossRef] [PubMed]

53. Vuilleumier, S.; Fiorentino, A.; Dénéréaz, S.; Spichiger, T. Identification of new demands regarding prehospital care based on 35,188 missions in 2018. *BMC Emerg. Med.* **2021**, *21*, 63. [CrossRef] [PubMed]
54. Andrew, E.; Nehme, Z.; Cameron, P.; Smith, K. Drivers of Increasing Emergency Ambulance Demand. *Prehospital Emerg. Care* **2020**, *24*, 385–393. [CrossRef] [PubMed]
55. NHS England. The NHS Long Term Plan. 2019. Available online: https://webarchive.nationalarchives.gov.uk/ukgwa/20250506042016/https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ (accessed on 28 July 2026).
56. Fry, A.; Littlejohns, T.J.; Sudlow, C.; Doherty, N.; Adamska, L.; Sprosen, T.; Collins, R.; Allen, N.E. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **2017**, *186*, 1026–1034. [CrossRef] [PubMed]
57. Otles, E.; Oh, J.; Li, B.; Bochinski, M.; Joo, H.; Ortwine, J.; Shenoy, E.; Washer, L.; Young, V.B.; Rao, K. Mind the performance gap: Examining dataset shift during prospective validation. *Proc. Mach. Learn. Res.* **2021**, *149*, 506–534. [PubMed]
58. Grisel, B.; Adisa, O.; Sakita, F.M.; Tarimo, T.G.; Kweka, G.L.; Mlangi, J.J.; Maro, A.V.; Yamamoto, M.; Coaxum, L.; Arthur, D. Evaluating the performance of the HEART score in a Tanzanian emergency department. *Acad. Emerg. Med.* **2024**, *31*, 361–370. [CrossRef] [PubMed]
59. Debray, T.P.A.; Collins, G.S.; Riley, R.D.; I E Snell, K.; Van Calster, B.; Reitsma, J.B.; Moons, K.G.M. Transparent reporting of multivariable prediction models developed or validated using clustered data: TRIPOD-Cluster checklist. *BMJ* **2023**, *380*, e071018. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.