

Resting Clinical and Summarized PET Variables for Prescreening Reduced Coronary Flow Reserve: A Leakage-Controlled Single-Center Feasibility Study

Ruonan Wang

The First Hospital of Shanxi Medical University, Shanxi Medical University

Yiting Liu

The First Hospital of Shanxi Medical University, Shanxi Medical University

Zitong Sang

The First Hospital of Shanxi Medical University, Shanxi Medical University

Haitao Zhou

The First Hospital of Shanxi Medical University, Shanxi Medical University

Zhifang Wu

The First Hospital of Shanxi Medical University, Shanxi Medical University

Sijin Li

sijinl@sxmu.edu.cn

The First Hospital of Shanxi Medical University, Shanxi Medical University

Research Article

Keywords: Cardiac PET, coronary flow reserve, risk stratification, prescreening, temporal validation, feasibility study

Posted Date: August 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10184863/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Coronary flow reserve (CFR) derived from stress PET provides clinically important physiologic information, but a practical question remains whether variables available at rest can support low-CFR prescreening in selected clinical settings. This question should be distinguished from using resting information as a substitute for standard stress-based physiologic assessment.

Methods

We performed a single-center retrospective feasibility study of 354 patients who underwent cardiac PET with quantifiable CFR. Candidate predictors were restricted to resting clinical variables, resting hemodynamics, and summarized gated/rest PET parameters. All stress variables, myocardial blood flow measurements, and CFR-derived variables were excluded before modeling to prevent information leakage. We evaluated continuous CFR prediction and binary classification for reduced CFR using repeated internal cross-validation. The primary prescreening endpoint was $\text{CFR} < 2.0$, and $\text{CFR} < 2.5$ was assessed as a secondary endpoint. To prioritize a stricter estimate of transportability, we additionally performed internal temporal validation using a fixed 2021-01-01 split, model-block comparison, and threshold-stability analyses based on bootstrap confidence intervals and repeated nested resampling.

Results

In internal temporal validation, which was treated as the primary analysis, discrimination for the primary endpoint was weak: test area under the receiver operating characteristic curve (AUC) was 0.515 for clinical variables alone, 0.589 after adding resting hemodynamics, and 0.599 after adding summarized imaging variables. For $\text{CFR} < 2.5$, the corresponding temporal test AUCs were 0.519, 0.608, and 0.657, indicating only limited discrimination. For $\text{CFR} < 2.0$, the apparent training-derived threshold targeting sensitivity ≥ 0.85 did not transport well, with temporal test sensitivity 0.200 and a bootstrap 95% interval of 0.000-0.500. For $\text{CFR} < 2.5$, temporal test sensitivity and specificity were 0.581 and 0.588, respectively. Repeated nested resampling showed broad threshold distributions and marked sensitivity-specificity trade-offs for both endpoints, indicating that uncertainty could be characterized more clearly but not converted into a stable screening cut-off. No stable fixed threshold was identified in the current dataset. Internal repeated cross-validation showed weak continuous CFR prediction (best out-of-fold performance: $R = 0.275$, $\text{MAE} = 0.640$, $\text{RMSE} = 0.908$) and only modest discrimination overall.

Conclusions

Resting clinical and summarized imaging variables did not provide performance consistent with a substitute for measured CFR or standard stress PET assessment. They may provide limited value for low-CFR prescreening or risk stratification, but the signal was weak for $\text{CFR} < 2.0$ and only modest for $\text{CFR} < 2.5$. No stable fixed threshold was identified in the current dataset. The main contribution of this study is therefore not a deployable rest-only model, but a leakage-controlled benchmark that clarifies the clinical boundary between rest-only prescreening and physiologic measurement.

Background

PET-derived myocardial blood flow and CFR provide clinically important information on coronary physiology, including risk related to both obstructive coronary disease and microvascular dysfunction [1–4]. In daily practice, however, obtaining this information requires stress acquisition and a full quantitative protocol, which creates a practical incentive to ask whether data available at rest might still provide useful low-CFR prescreening information [2, 5].

This prescreening question is clinically relevant, but it is methodologically different from attempting to reconstruct a physiologic reserve measure from baseline data alone. CFR reflects the circulatory response to vasodilator stress rather than resting status alone [5–11]. As a result, any study that asks a rest-only question must separate clinically available resting information from variables that are only known after stress acquisition or directly encode the target itself.

That distinction is particularly important in cardiac PET studies because apparently strong models can become misleading if they include stress hemodynamics, stress imaging variables, myocardial blood flow measurements, or CFR-derived quantities. Such inputs do not test a true rest-only hypothesis and may create substantial optimism through information leakage [13–17, 30]. For this reason, a negative or borderline result obtained under strict leakage control may be more informative than an apparently stronger result derived from a less constrained feature set.

Accordingly, the aim of the present study was not to maximize discrimination from all available PET-derived information, but to estimate the limited signal available from routine rest-only inputs under a methodologically strict rest-only design.

The present study therefore asked a narrower and more defensible question: can resting clinical variables and summarized gated/rest PET variables provide useful prescreening or risk-stratification information for reduced CFR in a single-center cohort? Our primary objective was to assess

performance for the clinically stricter endpoint of CFR < 2.0, with CFR < 2.5 analyzed as a secondary endpoint [4, 18, 29]. Our secondary objective was to determine whether such rest-only information can be interpreted only as limited prescreening support rather than as a substitute for measured CFR or standard stress PET assessment.

Methods

Study design and setting

This was a retrospective single-center feasibility study based on a cardiac PET cohort with quantifiable CFR. All analyzable cases available in the processed study dataset were screened. The analysis was restricted to prespecified rest-only variables in order to address a clearly defined methodological question under strict leakage control.

Cohort definition

The screened dataset contained 354 consecutive examinations with available CFR reference values. No cases were excluded for missing CFR, and the final analyzable cohort therefore also consisted of 354 patients. For the primary descriptive comparison, patients were grouped by CFR < 2.0 versus CFR $\geq$ 2.0.

Figure 1 summarizes cohort selection and predictor restriction.

Figure 1. Cohort selection and predictor restriction process Consecutive PET studies with available CFR (n=354) were screened, and all were retained in the final cohort. Predictor candidates were restricted to resting clinical, hemodynamic, and summarized gated/rest PET variables. All stress variables, myocardial blood flow measurements, and CFR-derived quantities were excluded before modeling to preserve a strict leakage-controlled rest-only question.

Targets

The continuous target was CFR. Binary prescreening endpoints were predefined as:

- primary endpoint: CFR < 2.0
- secondary endpoint: CFR < 2.5

These thresholds were chosen because they are commonly used in clinically oriented CFR interpretation and contemporary prognostic PET/CFR literature [4,18,29], while acknowledging that threshold-based discrimination does not by itself imply clinical deployment readiness.

Candidate predictors and leakage control

Candidate predictors were restricted to information that could reasonably be considered available at rest, including:

- demographics and anthropometric variables
- resting comorbidity indicators
- resting medication indicators
- resting systolic blood pressure, diastolic blood pressure, and heart rate
- summarized gated/rest PET functional variables

The following variables were excluded a priori before model development:

- all stress hemodynamic variables
- all stress imaging variables
- rest myocardial blood flow
- stress myocardial blood flow
- regional CFR variables
- any other CFR-derived quantities or direct target components

This restriction was used to preserve a leakage-controlled rest-only question rather than a mixed rest-plus-stress prediction setup.

Potentially informative stress, flow, or other target-proximal variables were therefore excluded by design, not because they lacked signal, but because their inclusion would have changed the study question.

Table 2 summarizes the allowable and excluded feature groups.

Table 2. Leakage-controlled feature groups and exclusions

Panel A. Allowed rest-only input groups

Feature group	Variables included in the prespecified analysis	Clinical role	Reason for inclusion
Demographics and anthropometrics	sex, age, height, weight, body mass index	baseline patient profile	routinely available before stress acquisition
Resting comorbidity indicators	hypertension, diabetes, dyslipidemia, family history, smoking, alcohol use	cardiovascular risk burden	clinically available at rest
Resting medication indicators	statin, antiplatelet, beta-blocker, ACEI/ARB, calcium-channel blocker, diuretic, nitrate	treatment context	clinically available at rest
Resting hemodynamics	resting systolic blood pressure, resting diastolic blood pressure, resting heart rate	baseline circulatory status	obtainable before stress acquisition
Summarized gated/rest PET functional variables	M-rest EDV, M-rest ESV, M-rest EF, C-rest EDV, C-rest ESV, C-rest EF, QGS-rest EDV, QGS-rest ESV, QGS-rest EF, rest mass (C), SRS	summarized resting cardiac structure and function	available from resting gated imaging without direct use of flow-reserve measurements

Panel B. Excluded variables

Excluded group	Examples	Reason for exclusion
Stress hemodynamics	stress SBP, stress DBP, stress HR, delta HR, HR reserve/recovery terms	unavailable in a true rest-only scenario
Stress imaging variables	stress EDV, stress ESV, stress EF, stress wall/perfusion summaries	acquired only after stress testing
Myocardial blood flow measurements	rest MBF, stress MBF	directly tied to flow quantification and would blur the rest-only prescreening question
CFR-derived variables	global CFR, regional CFR, CFR-minimum terms, related derived ratios	direct target leakage
Regional stress-flow descriptors	regional stress segment flow summaries and derived minimums	post-stress physiologic information not allowed in a strict rest-only analysis

ACEI indicates angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; EDV, end-diastolic volume; ESV, end-systolic volume; EF, ejection fraction; PET, positron emission tomography; SBP, systolic blood pressure; DBP, diastolic blood pressure; SRS, summed rest score. This table defines the prespecified feature policy applied before all modeling steps.

Data preparation

The analysis dataset was derived from the cleaned source spreadsheet using a predefined variable schema and leakage exclusion rules. Missing values were handled inside the model pipelines during validation rather than by complete-case restriction, because the latter would have reduced the usable cohort to an implausibly small subset. Variable audit, missingness review, duplicate checks, and outlier screening were completed before modeling.

Modeling strategy

For continuous CFR prediction, we used:

- ridge regression as the main interpretable baseline model
- random forest regression as a predefined nonlinear comparator

For binary low-CFR classification, we used:

- logistic regression as the main prespecified classifier
- random forest classifier as a predefined nonlinear comparator

The analysis was intentionally limited to prespecified interpretable and nonlinear comparators in order to preserve a clinically bounded feasibility framework rather than pursue marginal algorithmic gains.

Validation strategy

We used two validation layers.

First, repeated internal cross-validation was used to estimate overall internal performance stability while preserving preprocessing within each validation loop. These analyses were treated as supportive rather than definitive.

Second, and more importantly for study interpretation, we performed internal temporal validation using a fixed split at 2021-01-01. Cases acquired before this date formed the training set, and later cases formed the temporal test set. This split yielded 255 training cases and 99 temporal test cases, including 10 primary-endpoint events (CFR < 2.0) and 31 secondary-endpoint events (CFR < 2.5) in the test set. This temporal split was prespecified and was not iteratively optimized against model performance.

Model-block comparison

To clarify where any predictive signal arose, we prespecified three logistic-regression input blocks for binary analyses:

1. clinical only
2. clinical plus resting hemodynamics
3. clinical plus resting hemodynamics plus summarized imaging variables

This comparison was used to evaluate whether resting hemodynamics and summarized imaging added clinically meaningful incremental discrimination beyond routine clinical information.

Threshold transport and stability analysis

For each binary endpoint, we first performed the prespecified threshold-transport check: a threshold was selected on the temporal-training set to achieve sensitivity ≥ 0.85 whenever possible and then transferred unchanged to the temporal test set. We reported sensitivity, specificity, positive predictive value, negative predictive value, and rule-out rate.

To characterize threshold uncertainty rather than improve apparent performance, we additionally performed threshold-stability analyses using stratified bootstrap confidence intervals for temporal-test threshold metrics and repeated nested resampling. In these analyses, thresholds were selected from out-of-fold predictions within the training portion and then evaluated on held-out data. This allowed assessment of threshold dispersion, sensitivity-specificity trade-offs, and whether any fixed cut-off remained stable under repeated internal perturbation.

Outcomes and performance metrics

For continuous CFR prediction, we reported Pearson correlation (R), mean absolute error (MAE), root mean squared error (RMSE), and calibration slope/intercept.

For binary classification, we reported area under the receiver operating characteristic curve (AUC), sensitivity, specificity, balanced accuracy, calibration slope/intercept, and for threshold transport and stability analyses also positive predictive value, negative predictive value, and rule-out rate.

Reporting framework

This study is reported as a prediction-oriented observational feasibility study with explicit leakage control and temporal validation, using TRIPOD-oriented principles for transparent multivariable prediction reporting and explicit description of the validation structure and study limitations [13-14].

Results

Cohort characteristics

The analyzable cohort included 354 patients. The primary endpoint CFR < 2.0 was present in 33 cases, and the secondary endpoint CFR < 2.5 was present in 93 cases. Baseline characteristics for the full cohort and for the primary endpoint groups are shown in Table 1. Patients with CFR < 2.0 had descriptively higher resting systolic blood pressure and heart rate and a higher prevalence of hypertension, but the purpose of Table 1 is cohort characterization rather than univariable claim making.

Table 1. Baseline characteristics by primary CFR category

Characteristic	Overall (N=354)	CFR < 2.0 (n=33)	CFR ≥ 2.0 (n=321)	Missing, n
Age, years	53.0 [47.0, 59.0]	55.0 [49.8, 60.0]	53.0 [46.5, 59.0]	7
BMI, kg/m ²	25.4 [22.9, 27.6]	26.4 [24.3, 29.1]	25.2 [22.9, 27.3]	11
Rest SBP, mmHg	126.0 [113.0, 138.8]	136.5 [123.5, 145.0]	126.0 [112.2, 138.0]	24
Rest DBP, mmHg	72.0 [64.0, 80.0]	74.5 [65.8, 81.0]	71.0 [64.0, 80.0]	24
Rest HR, bpm	65.0 [60.0, 72.0]	72.0 [63.0, 79.0]	65.0 [59.8, 71.2]	21
M-rest EF, %	59.0 [54.0, 63.0]	63.0 [56.5, 67.0]	58.0 [53.0, 62.0]	29
C-rest EF, %	55.0 [52.0, 59.0]	59.0 [55.0, 62.0]	55.0 [52.0, 58.0]	50
QGS-rest EF, %	50.0 [44.8, 56.0]	56.0 [51.0, 60.0]	50.0 [44.0, 55.0]	66
Rest mass (C), g	73.0 [62.0, 86.2]	69.0 [60.0, 83.0]	75.0 [62.0, 87.0]	50
SRS	1.0 [0.0, 2.0]	1.0 [0.0, 3.0]	1.0 [0.0, 2.0]	68
Female sex, n (%)	149/354 (42.1)	11/33 (33.3)	138/321 (43.0)	0
Hypertension, n (%)	149/320 (46.6)	23/31 (74.2)	126/289 (43.6)	34
Diabetes, n (%)	75/320 (23.4)	6/31 (19.4)	69/289 (23.9)	34
Dyslipidemia, n (%)	161/314 (51.3)	16/30 (53.3)	145/284 (51.1)	40
Family history, n (%)	104/319 (32.6)	7/30 (23.3)	97/289 (33.6)	35
Current smoking, n (%)	117/309 (37.9)	12/29 (41.4)	105/280 (37.5)	45
Alcohol use, n (%)	75/307 (24.4)	10/29 (34.5)	65/278 (23.4)	47
Statin use, n (%)	156/284 (54.9)	15/27 (55.6)	141/257 (54.9)	70
Antiplatelet use, n (%)	115/283 (40.6)	8/26 (30.8)	107/257 (41.6)	71
Beta-blocker use, n (%)	71/284 (25.0)	9/27 (33.3)	62/257 (24.1)	70
ACEI/ARB use, n (%)	83/282 (29.4)	10/27 (37.0)	73/255 (28.6)	72
Calcium-channel blocker use, n (%)	78/287 (27.2)	14/28 (50.0)	64/259 (24.7)	67
Diuretic use, n (%)	13/284 (4.6)	3/27 (11.1)	10/257 (3.9)	70
Nitrate use, n (%)	35/282 (12.4)	4/26 (15.4)	31/256 (12.1)	72

Continuous variables are presented as median [interquartile range]. Categorical variables are presented as n/N (%), using the number of non-missing observations for each variable as the denominator. CFR < 2.0 was the primary low-CFR endpoint for prescreening analyses. Baseline differences are shown descriptively to characterize the cohort and were not used to claim causal or diagnostic separation.

Table 1 summarizes baseline characteristics of the analyzable cohort.

Internal temporal validation

Because the clinical question concerns whether a rest-only signal can persist beyond an internally optimistic setting, the temporal split results are the primary results of the study.

For the primary endpoint CFR < 2.0, temporal discrimination was weak. Using logistic regression, the temporal test AUC was 0.515 for clinical variables alone, 0.589 after adding resting hemodynamics, and 0.599 after adding summarized imaging variables. These findings indicate that routine resting clinical variables alone were close to non-informative for the stricter endpoint and that the addition of resting hemodynamics and summarized imaging variables provided only modest gain.

For the secondary endpoint CFR < 2.5, the corresponding temporal test AUCs were 0.519, 0.608, and 0.657. This pattern suggests that resting hemodynamics and summarized imaging variables do contain some information related to less stringent low-CFR categorization, but the magnitude of discrimination remains limited.

Table 3 summarizes model performance across internal and temporal validation.

Table 3. Internal and temporal model performance summary

Panel A. Continuous CFR prediction under repeated internal cross-validation

Model	Validation setting	R (95% interval)	MAE (95% interval)	RMSE (95% interval)	Calibration slope	Calibration intercept
Ridge regression	Repeated internal cross-validation	0.272 (0.244 to 0.277)	0.651 (0.647 to 0.657)	0.909 (0.908 to 0.916)	1.045	-0.142
Random forest regression	Repeated internal cross-validation	0.275 (0.233 to 0.300)	0.640 (0.633 to 0.649)	0.908 (0.901 to 0.920)	1.023	-0.080

Panel B. Binary low-CFR classification under repeated internal cross-validation

Endpoint	Model	AUC (95% interval)	Sensitivity	Specificity	Balanced accuracy (95% interval)	Calibration slope	Calibration intercept
CFR < 2.0	Logistic regression	0.703 (0.640 to 0.723)	0.000	1.000	0.500 (0.499 to 0.507)	1.161	0.375
CFR < 2.0	Random forest classifier	0.688 (0.652 to 0.702)	0.000	0.978	0.489 (0.491 to 0.538)	0.766	-1.231
CFR < 2.5	Logistic regression	0.677 (0.648 to 0.678)	0.054	0.966	0.510 (0.507 to 0.532)	0.951	-0.043
CFR < 2.5	Random forest classifier	0.645 (0.625 to 0.658)	0.323	0.874	0.598 (0.556 to 0.613)	0.866	-0.572

Panel C. Binary prescreening performance under internal temporal validation

Endpoint	Feature block	Number of features	Temporal test AUC	Temporal test Brier score	Temporal calibration slope	Temporal calibration intercept
CFR < 2.0	Clinical only	18	0.515	0.105	0.061	-1.968
CFR < 2.0	Clinical + resting hemodynamics	21	0.589	0.092	0.375	-1.201
CFR < 2.0	Clinical + resting hemodynamics + summarized imaging	32	0.599	0.095	0.285	-1.301
CFR < 2.5	Clinical only	18	0.519	0.220	0.293	-0.449
CFR < 2.5	Clinical + resting hemodynamics	21	0.608	0.210	0.727	0.162
CFR < 2.5	Clinical + resting hemodynamics + summarized imaging	32	0.657	0.208	1.294	0.773

Continuous-task intervals and repeated cross-validation intervals are empirical 2.5th to 97.5th percentile ranges across repeated validation runs in the prespecified analysis set. Temporal validation results are single-split point estimates from an internal temporal split at 2021-01-01 and therefore are reported without additional confidence intervals. Temporal validation was the primary validation layer for study interpretation.

Figure 2 summarizes temporal discrimination across prespecified model blocks.

Figure 2. Temporal model-block discrimination Temporal test AUCs are shown for three prespecified logistic-regression input blocks: clinical only, clinical plus resting hemodynamics, and clinical plus resting hemodynamics plus summarized imaging variables. Results are shown separately for the primary endpoint (CFR < 2.0) and the secondary endpoint (CFR < 2.5).

Threshold transport and stability analysis

For CFR < 2.0, the apparent temporal-training threshold selected to target sensitivity ≥ 0.85 was 0.0818. In the training set, this produced sensitivity 0.870, specificity 0.694, negative predictive value 0.982, and a rule-out rate of 0.643. When the same threshold was transferred unchanged to the temporal test set, however, sensitivity dropped to 0.200, while specificity was 0.809, negative predictive value was 0.900, and the rule-out rate was 0.808. The bootstrap 95% interval for temporal-test sensitivity was 0.000-0.500, indicating substantial uncertainty around this apparent operating point.

When threshold selection was based on out-of-fold training predictions rather than apparent training predictions, the temporal threshold for CFR < 2.0 shifted to 0.0519, yielding temporal-test sensitivity 0.400 and specificity 0.629. Repeated nested resampling showed that threshold values remained broadly dispersed (median 0.0651, 95% range 0.0054-0.0855) and that threshold-dependent performance varied substantially, with outer-test sensitivity 0.929 [0.480-1.000] and outer-test specificity 0.312 [0.000-0.598]. These patterns indicate a pronounced sensitivity-specificity trade-off rather than identification of a stable fixed threshold for the primary endpoint.

For CFR < 2.5, the apparent temporal-training threshold was 0.2224. In the temporal test set, sensitivity and specificity were 0.581 and 0.588, respectively, with a negative predictive value of 0.755. The bootstrap 95% interval for temporal-test sensitivity was 0.419-0.742. When threshold selection was based on out-of-fold training predictions, the temporal threshold shifted to 0.1946, yielding sensitivity 0.871 and specificity 0.382, again reflecting movement along the same sensitivity-specificity trade-off rather than improved transportability. In repeated nested resampling, threshold values remained variable (median 0.1806, 95% range 0.1205-0.2163), while outer-test sensitivity and specificity were 0.895 [0.632-1.000] and 0.308 [0.033-0.525], respectively. These findings support only limited prescreening value and do not identify a stable fixed threshold for the secondary endpoint.

Taken together, the more robust threshold-stability analyses quantified uncertainty more clearly but did not resolve the transport problem observed on temporal validation. No stable fixed threshold was identified in the current dataset.

Table 4 summarizes prespecified threshold transport metrics.

Table 4. Prespecified threshold transport metrics

Endpoint	Dataset	Threshold selected on training set	Target training sensitivity	AUC	Brier score	Sensitivity	Specificity	PPV	NPV	Rule-out rate	TN	FP	FN	TP
CFR < 2.0	Training set	0.082	0.850	0.882	0.062	0.870	0.694	0.220	0.982	0.643	161	71	3	20
CFR < 2.0	Temporal test set	0.082	0.850	0.599	0.095	0.200	0.809	0.105	0.900	0.808	72	17	8	2
CFR < 2.5	Training set	0.222	0.850	0.762	0.164	0.855	0.534	0.371	0.920	0.439	103	90	9	53
CFR < 2.5	Temporal test set	0.222	0.850	0.657	0.208	0.581	0.588	0.391	0.755	0.535	40	28	13	18

Thresholds were selected on the temporal-training set using the full prespecified logistic-regression model with clinical, resting hemodynamic, and summarized imaging variables, targeting sensitivity ≥ 0.85 whenever possible. The same threshold was then transferred unchanged to the internal temporal test set. These values should be interpreted as apparent threshold-transport checks rather than as recommended operational cut-offs. CFR < 2.0 was the primary endpoint and CFR < 2.5 was the secondary endpoint. PPV indicates positive predictive value; NPV, negative predictive value.

Figure 3 shows temporal calibration by endpoint.

Figure 3. Temporal calibration by endpoint Calibration curves on the temporal test set for the full prespecified logistic-regression model are shown separately for CFR < 2.0 and CFR < 2.5. The diagonal line indicates ideal calibration. These plots are intended to complement the threshold transport and stability analyses rather than to imply external validation.

Internal cross-validation results

Repeated internal cross-validation was performed as a supportive analysis. For continuous CFR prediction, the best out-of-fold performance was observed with random forest regression (R=0.275, MAE=0.640, RMSE=0.908), while ridge regression performed similarly (R=0.272, MAE=0.651,

RMSE=0.909). These values indicate weak continuous prediction overall.

For binary classification, the best out-of-fold AUC for the primary endpoint CFR < 2.0 was 0.703 with logistic regression, whereas the best out-of-fold AUC for CFR < 2.5 was 0.677, also with logistic regression. Although these internal values suggest weak-to-modest ranking signal, they should be interpreted cautiously and not as evidence of deployment readiness, because the stricter temporal validation results were less favorable, especially for the primary endpoint.

Figure 4 provides descriptive CFR distributions for the cohort and prespecified subgroups.

Figure 4. CFR distribution and subgroup overview Panel A shows the distribution of measured CFR in the full analyzable cohort. Panel B summarizes CFR distributions across prespecified clinical subgroups, including sex, age, oCAD status, hypertension, and diabetes. These descriptive plots provide cohort context and are not intended to imply subgroup-specific model validity.

Summary of the incremental signal structure

Across analyses, the signal structure was consistent. Routine clinical variables alone were insufficient for useful discrimination. Resting hemodynamics added some information, and summarized imaging variables added a further but modest increment, particularly for CFR < 2.5. However, these additions did not produce a level of temporal performance that would support stronger clinical claims.

Discussion

Principal findings

This study addressed a deliberately strict rest-only question and found that resting clinical variables and summarized gated PET variables provided, at most, limited low-CFR prescreening information. The strongest and most clinically relevant negative finding was for the primary endpoint CFR < 2.0: temporal discrimination remained weak, a training-derived high-sensitivity threshold failed to maintain sensitivity on later cases, and no stable fixed threshold was identified after more robust reassessment of threshold uncertainty. For CFR < 2.5, the signal was somewhat stronger, but still remained in the range of limited discrimination rather than a robust clinically actionable result. The study should therefore be read primarily as a boundary-setting feasibility analysis rather than as a near-deployable screening model.

Why the rest-only signal was limited

The biological explanation is plausible. CFR is a dynamic measure of physiologic reserve rather than a static resting descriptor [5–11, 19]. Resting clinical characteristics, baseline hemodynamics, and summarized functional imaging variables can reflect underlying burden of disease or cardiovascular status, but they do not fully capture vasodilator response. It is therefore unsurprising that the strongest rest-only signals in this dataset were modest and that the stricter endpoint deteriorated under temporal validation.

Why strict leakage control matters

An important contribution of this study is methodological. In cardiac PET research, the apparent distinction between prediction and direct target encoding can become blurred if investigators mix true resting information with stress or flow-derived variables. The present analysis excluded all such variables before modeling, thereby preserving a methodologically coherent rest-only question. Under that constraint, the results define a more credible lower-bound benchmark for what can realistically be expected from routine resting inputs [13–17, 30]. In this sense, the negative or borderline findings are not incidental; they are the substantive result of a deliberately constrained and clinically interpretable design.

Interpretation of the model-block comparison

The model-block comparison provides a useful structural interpretation of the signal. Routine clinical variables alone contributed little discriminatory value. Resting hemodynamics improved performance modestly, suggesting that some low-CFR signal is present in baseline circulatory status. Summarized imaging variables added a further increment, especially for CFR < 2.5, indicating that resting functional imaging still captures some downstream manifestation of disease burden. However, this incremental signal remained insufficient to support stronger claims about broader physiologic assessment.

Clinical interpretation

The appropriate clinical interpretation is narrow. A rest-only model of this kind may help frame exploratory prescreening or research-level risk stratification for reduced CFR, particularly for less stringent low-CFR definitions. However, the present results do not support using routine resting variables as a stand-alone basis for withholding or deferring indicated stress PET-based physiologic assessment. In particular, the poor temporal transport of the apparent high-sensitivity threshold for CFR < 2.0, together with the absence of any stable fixed threshold under repeated nested resampling, argues against any stronger clinical positioning. Even for CFR < 2.5, threshold-dependent performance remained dominated by sensitivity-specificity trade-offs rather than by identification of a robust screening cut-off.

Interpretation of threshold instability

A key methodological finding of this study is that threshold-dependent performance was not stable. For $\text{CFR} < 2.0$, the temporal test set contained only 10 events, so each additional misclassified event changed sensitivity by 0.10. This alone made high-sensitivity threshold estimates highly discrete and uncertain. In addition, selecting a threshold on fitted temporal-training predictions introduced optimism, which was reflected by the gap between apparent training performance and later temporal-test performance.

The instability was not explained by a single unlucky temporal split alone. Repeated nested resampling showed broad distributions of selected thresholds and wide dispersions of threshold-dependent sensitivity and specificity for both endpoints, indicating that the problem persisted under repeated internal perturbation. Temporal drift also appeared plausible: several rest-only predictors, including statin use, antiplatelet use, family history, dyslipidemia, and sex distribution, differed meaningfully between temporal subsets. Finally, the underlying classification signal remained modest under strict temporal validation, especially for $\text{CFR} < 2.0$. Taken together, these features explain why a threshold that appeared favorable in one subset did not support a stable prescreening cut-off.

Relation to existing literature

Prior PET/CFR literature has established the clinical relevance of measured CFR and the importance of stress-provoked physiology [1–4, 12, 18, 20–26]. The present study does not challenge that literature. Instead, it addresses a narrower translational question and shows that once stress and flow-derived leakage sources are excluded, the remaining rest-only signal is limited. This distinction is important because optimism can arise when richer but methodologically mixed inputs are interpreted as if they represented a true rest-only approach. Related attempts to infer hyperemic physiology from reduced input spaces have also been explored in invasive settings using resting pressure curves or simplified reconstruction-based calculations [27–28], further underscoring how difficult it is to derive reserve physiology without direct hyperemic measurement.

Limitations

This study has several limitations. First, it was a retrospective single-center analysis with a limited number of low-CFR events, particularly for the primary endpoint of $\text{CFR} < 2.0$, which constrained model stability and precision in light of general prediction-model sample size considerations [15–16]. The same constraint also limited threshold stability, because threshold-dependent sensitivity for $\text{CFR} < 2.0$ was estimated from only 10 temporal-test events. Second, only routinely available resting clinical variables and summarized gated/rest imaging parameters were included; although this restriction was deliberate to preserve a true rest-only question and avoid leakage, it also means that potentially informative higher-resolution image data, structured electrocardiographic features, and more complete laboratory or angiographic descriptors were not evaluated. Third, the temporal split used here provides a stricter test than random internal validation but does not replace true external validation across centers, scanners, acquisition protocols, and processing conditions [14]. Fourth, several candidate variables showed missingness or heterogeneity typical of real-world retrospective data, and modest temporal drift was present in several rest-only predictors. Finally, because the study was designed specifically to exclude stress variables, myocardial blood flow measurements, and CFR-derived quantities, the results should be interpreted as a conservative benchmark for rest-only prescreening rather than as an estimate of the best possible performance achievable with richer multimodal data [13, 17]. Accordingly, the present threshold analyses are better interpreted as quantifying uncertainty and transport instability than as supporting deployment of a fixed prescreening cut-off.

Follow-up major adverse cardiovascular event data were also considered for inclusion, but were not incorporated into the formal analyses. This was because the endpoint composition differed between the two follow-up rounds, the number of cases with clearly adjudicable outcome status was limited, cancer-related special coding introduced additional complexity, some event records required manual adjudication, and the available time information was insufficient to support a robust time-to-event analysis.

Conclusions

In this single-center feasibility study, resting clinical and summarized imaging variables did not provide performance consistent with a substitute for measured CFR or standard stress PET assessment. They may provide limited value for low-CFR prescreening or risk stratification, but this value was weak for $\text{CFR} < 2.0$ and only modest for $\text{CFR} < 2.5$. No stable fixed threshold was identified in the current dataset. The most defensible interpretation of the present work is therefore not implementation readiness, but a leakage-controlled benchmark that clarifies the clinical boundary of rest-only prediction in cardiac PET.

Abbreviations

- AUC
area under the receiver operating characteristic curve
- CFR
coronary flow reserve
- DBP
diastolic blood pressure
- EF

ejection fraction

- HR

heart rate

- MAE

mean absolute error

- PET

positron emission tomography

- RMSE

root mean squared error

- SBP

systolic blood pressure

- SRS

summed rest score

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the First Hospital of Shanxi Medical University (Approval No. 2019-054). All procedures involving human participants were performed in accordance with the ethical standards of the institutional research committee, the Declaration of Helsinki, and relevant guidelines and regulations. The requirement for informed consent was waived by the Ethics Committee because of the retrospective nature of the study and the use of de-identified data.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' information

Not applicable.

Funding

This work was financially supported by the National Natural Science Foundation of China (Grant No. 62005150, U22A6008 and 82027804), Shanxi Province Higher Education Billion Project Science and Technology Guidance Project, Natural Science Foundation for Young Scientists of Shanxi Province (No. 202303021222344 and 202503021211260), and the First Hospital of Shanxi Medical University Hospital Fund Project (Project No. SYYYRC-2023004).

Author Contribution

R.W. conceived and designed the study, curated the data, performed the statistical analyses, prepared the figures, and drafted the manuscript. Y.L. and Z.S. participated in data collection and curation, contributed to the investigation, and revised the manuscript. H.Z. contributed to the study methodology, supervised the research, and revised the manuscript. Z.W. supervised the study, contributed to project administration, and revised the manuscript. S.L. conceived and supervised the study, acquired funding, contributed to project administration, and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available because they contain potentially identifiable patient information. The datasets are available from the corresponding author on reasonable request and with permission from the institutional ethics committee.

References

1. Murthy VL, Naya M, Foster CR, Hainer J, Gaber M, Di Carli G, et al. Improved cardiac risk assessment with noninvasive measures of coronary flow reserve. *Circulation*. 2011;124(20):2215–24. 10.1161/CIRCULATIONAHA.111.050427.
2. Murthy VL, Bateman TM, Beanlands RS, Berman DS, Borges-Neto S, Chareonthaitawee P, et al. Clinical quantification of myocardial blood flow using PET: joint position paper of the SNMMI Cardiovascular Council and the ASNC. *J Nucl Cardiol*. 2018;25(1):269–97. 10.1007/s12350-017-1110-x.
3. Green R, Cantoni V, Acampa W, Assante R, Zampella E, Nappi C, et al. Prognostic value of coronary flow reserve in patients with suspected or known coronary artery disease referred to PET myocardial perfusion imaging: a meta-analysis. *J Nucl Cardiol*. 2021;28(3):904–18. 10.1007/s12350-019-02000-7.
4. Schindler TH, Fearon WF, Pelletier-Galarneau M, Ambrosio G, Sechtem U, Ruddy TD, et al. Myocardial perfusion PET for the detection and reporting of coronary microvascular dysfunction: a JACC: Cardiovascular Imaging expert panel statement. *JACC Cardiovasc Imaging*. 2023;16(4):536–48. 10.1016/j.jcmg.2022.12.015.
5. Camici PG, Crea F. Coronary microvascular dysfunction. *N Engl J Med*. 2007;356(8):830–40. 10.1056/NEJMra061889.
6. Crea F, Camici PG, Bairey Merz CN. Coronary microvascular dysfunction: an update. *Eur Heart J*. 2014;35(17):1101–11. 10.1093/eurheartj/ehs513.
7. Camici PG, d'Amati G, Rimoldi O. Coronary microvascular dysfunction: mechanisms and functional assessment. *Nat Rev Cardiol*. 2015;12(1):48–62. 10.1038/nrcardio.2014.160.
8. Herrmann J, Kaski JC, Lerman A. Coronary microvascular dysfunction in the clinical setting: from mystery to reality. *Eur Heart J*. 2012;33(22):2771–83. 10.1093/eurheartj/ehs246.
9. Kaski JC, Crea F, Gersh BJ, Camici PG. Reappraisal of ischemic heart disease. *Circulation*. 2018;138(14):1463–80. 10.1161/CIRCULATIONAHA.118.031373.
10. Taqueti VR, Di Carli MF. Coronary microvascular disease pathogenic mechanisms and therapeutic options: JACC state-of-the-art review. *J Am Coll Cardiol*. 2018;72(21):2625–41. 10.1016/j.jacc.2018.09.042.
11. Tonet E, Pompei G, Faragasso E, Cossu A, Pavasini R, Passarini G, et al. Coronary microvascular dysfunction: PET, CMR and CT assessment. *J Clin Med*. 2021;10(9):1848. 10.3390/jcm10091848.
12. Ziadi MC. Myocardial flow reserve (MFR) with positron emission tomography/computed tomography: clinical impact in diagnosis and prognosis. *Cardiovasc Diagn Ther*. 2017;7(2):206–18. 10.21037/cdt.2017.04.10.
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594. 10.1136/bmj.g7594.
14. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol*. 2016;69:245–7. 10.1016/j.jclinepi.2015.04.005.
15. Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. 10.1136/bmj.m441.
16. Varoquaux G. Cross-validation failure: small sample sizes lead to large error bars. *NeuroImage*. 2018;180(Pt A):68–77. 10.1016/j.neuroimage.2017.06.061.
17. Varoquaux G, Cheplygina V. Machine learning for medical imaging: methodological failures and recommendations for the future. *NPJ Digit Med*. 2022;5:48. 10.1038/s41746-022-00592-y.
18. Murthy VL, Naya M, Taqueti VR, Foster CR, Gaber M, Hainer J, et al. Effects of sex on coronary microvascular dysfunction and cardiac outcomes. *Circulation*. 2014;129(24):2518–27. 10.1161/CIRCULATIONAHA.113.008507.
19. Johnson NP, Gould KL. How do PET myocardial blood flow reserve and FFR differ? *Curr Cardiol Rep*. 2020;22(4):20. 10.1007/s11886-020-1274-x.
20. Murthy VL, Naya M, Foster CR, Gaber M, Hainer J, Klein J, et al. Association between coronary vascular dysfunction and cardiac mortality in patients with and without diabetes mellitus. *Circulation*. 2012;126(15):1858–68. 10.1161/CIRCULATIONAHA.112.120402.
21. Taqueti VR, Solomon SD, Shah AM, Desai AS, Groarke JD, Osborne MT, et al. Coronary microvascular dysfunction and future risk of heart failure with preserved ejection fraction. *Eur Heart J*. 2018;39(10):840–9. 10.1093/eurheartj/ehx721.
22. Konerman MC, Greenberg JC, Koliass TJ, Corbett JR, Kritzman JN, Myerson M, et al. Reduced myocardial flow reserve is associated with diastolic dysfunction and decreased left atrial strain in patients with normal ejection fraction and epicardial perfusion. *J Card Fail*. 2018;24(2):90–100. 10.1016/j.cardfail.2017.10.010.
23. Valenta I, Dilsizian V, Quercioli A, Ruddy TD, Schindler TH. Quantitative PET/CT measures of myocardial flow reserve and atherosclerosis for cardiac risk assessment and predicting adverse patient outcomes. *Curr Cardiol Rep*. 2013;15(3):344. 10.1007/s11886-012-0344-0.
24. Gould KL, Kitkungvan D, Johnson NP, Nguyen T, Kirkeeide R, Bui L, et al. Mortality prediction by quantitative PET perfusion expressed as coronary flow capacity with and without revascularization. *JACC Cardiovasc Imaging*. 2021;14(5):1020–34. 10.1016/j.jcmg.2020.08.040.

25. de Winter RW, Jukema RA, van Diemen PA, Schumacher SP, van de Somsen YBO, et al. Prognostic value of modified coronary flow capacity derived from [15O]H₂O positron emission tomography perfusion imaging. *Circ Cardiovasc Imaging*. 2023;16(9):e014845. 10.1161/CIRCIMAGING.122.014845.
26. Naya M, Murthy VL, Foster CR, Hainer J, Gaber M, Klein J, et al. Preserved coronary flow reserve effectively excludes high-risk coronary artery disease on angiography. *J Nucl Med*. 2014;55(2):248–55. 10.2967/jnumed.113.121442.
27. Zimmermann F, Mast T, Johnson NP, Everts I, van Nunen L, Tonino P, et al. Deep learning for prediction of fractional flow reserve from resting coronary pressure curves. *EuroIntervention*. 2021;17(1):51–8. 10.4244/EIJ-D-20-00648.
28. Csippa B, Uveges A, Gyurki D, Jenei C, Tar B, Bugarin-Horvath B, et al. Simplified coronary flow reserve calculations based on three-dimensional coronary reconstruction and intracoronary pressure data. *Cardiol J*. 2023;30(4):516–25. 10.5603/CJ.a2021.0117.
29. Jensen SM, Prescott EIB, Abdulla J. The prognostic value of coronary flow reserve in patients with non-obstructive coronary artery disease and microvascular dysfunction: a systematic review and meta-analysis with focus on imaging modality and sex difference. *Int J Cardiovasc Imaging*. 2023;39(12):2545–56. 10.1007/s10554-023-02948-1.
30. Rosenblatt M, Tejavibulya L, Jiang R, Noble S, Scheinost D. Data leakage inflates prediction performance in connectome-based machine learning models. *Nat Commun*. 2024;15(1):1829. 10.1038/s41467-024-46150-w.

Figures

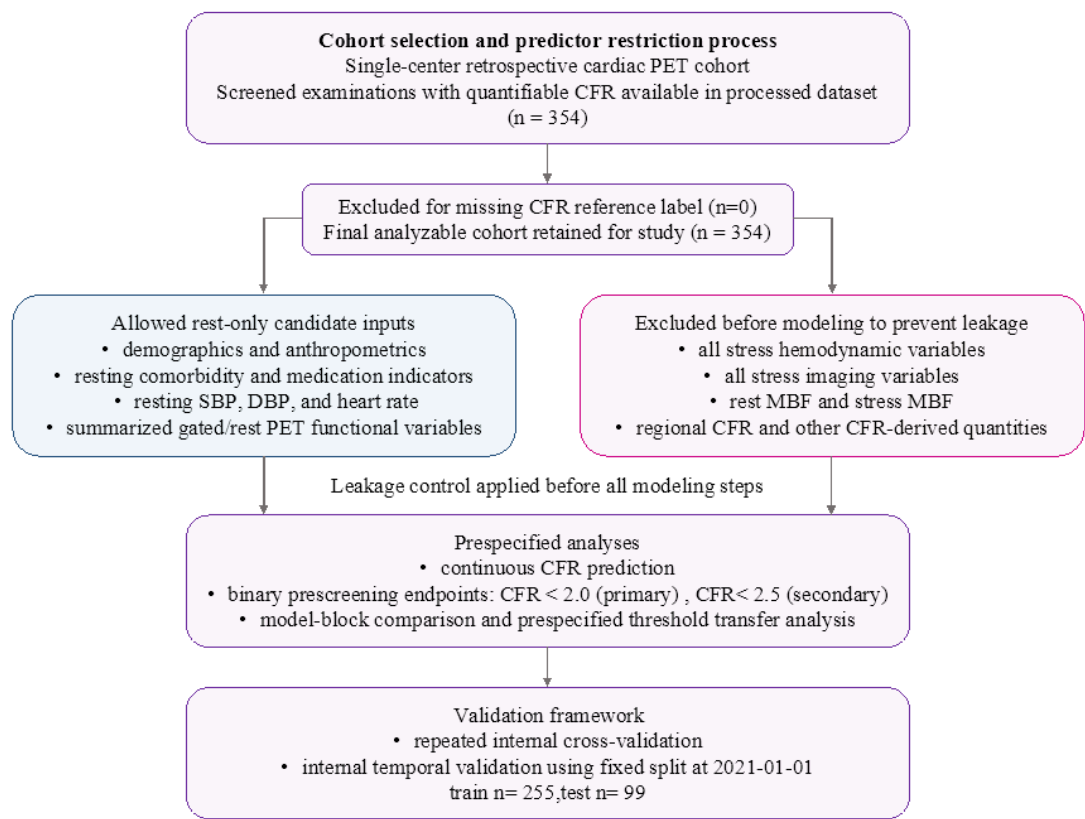


Figure 1

Cohort selection and predictor restriction process

Consecutive PET studies with available CFR (n=354) were screened, and all were retained in the final cohort. Predictor candidates were restricted to resting clinical, hemodynamic, and summarized gated/rest PET variables. All stress variables, myocardial blood flow measurements, and CFR-derived quantities were excluded before modeling to preserve a strict leakage-controlled rest-only question.

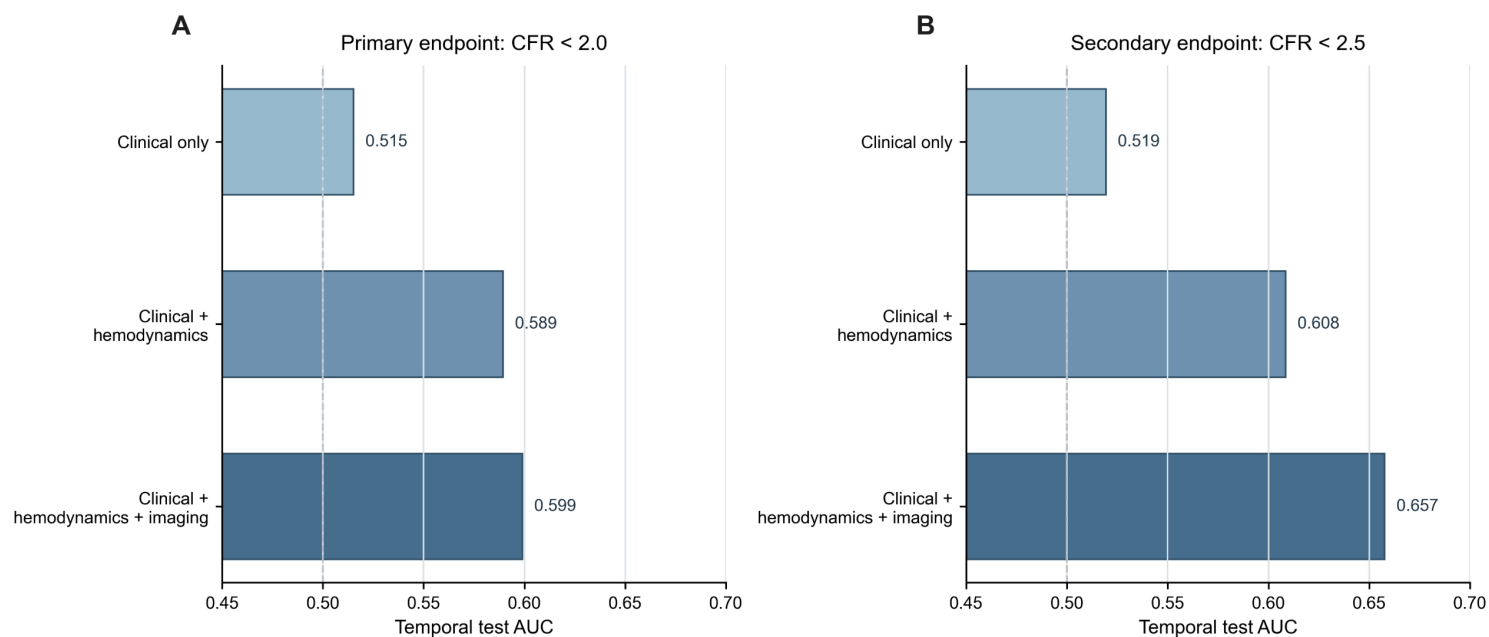


Figure 2

Temporal model-block discrimination

Temporal test AUCs are shown for three prespecified logistic-regression input blocks: clinical only, clinical plus resting hemodynamics, and clinical plus resting hemodynamics plus summarized imaging variables. Results are shown separately for the primary endpoint (CFR < 2.0) and the secondary endpoint (CFR < 2.5).

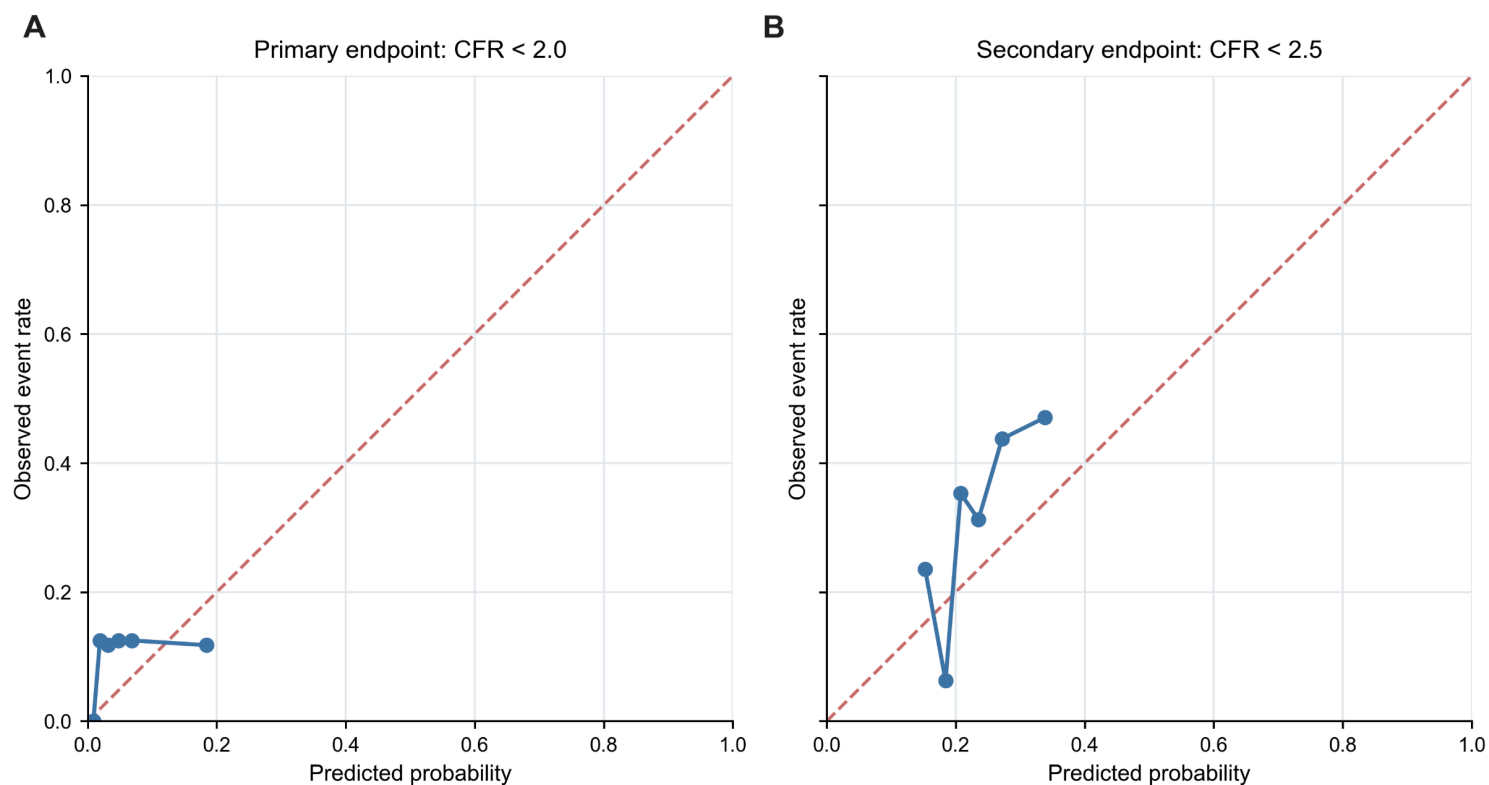


Figure 3

Temporal calibration by endpoint

Calibration curves on the temporal test set for the full prespecified logistic-regression model are shown separately for CFR < 2.0 and CFR < 2.5. The diagonal line indicates ideal calibration. These plots are intended to complement the threshold transport and stability analyses rather than to imply external validation.

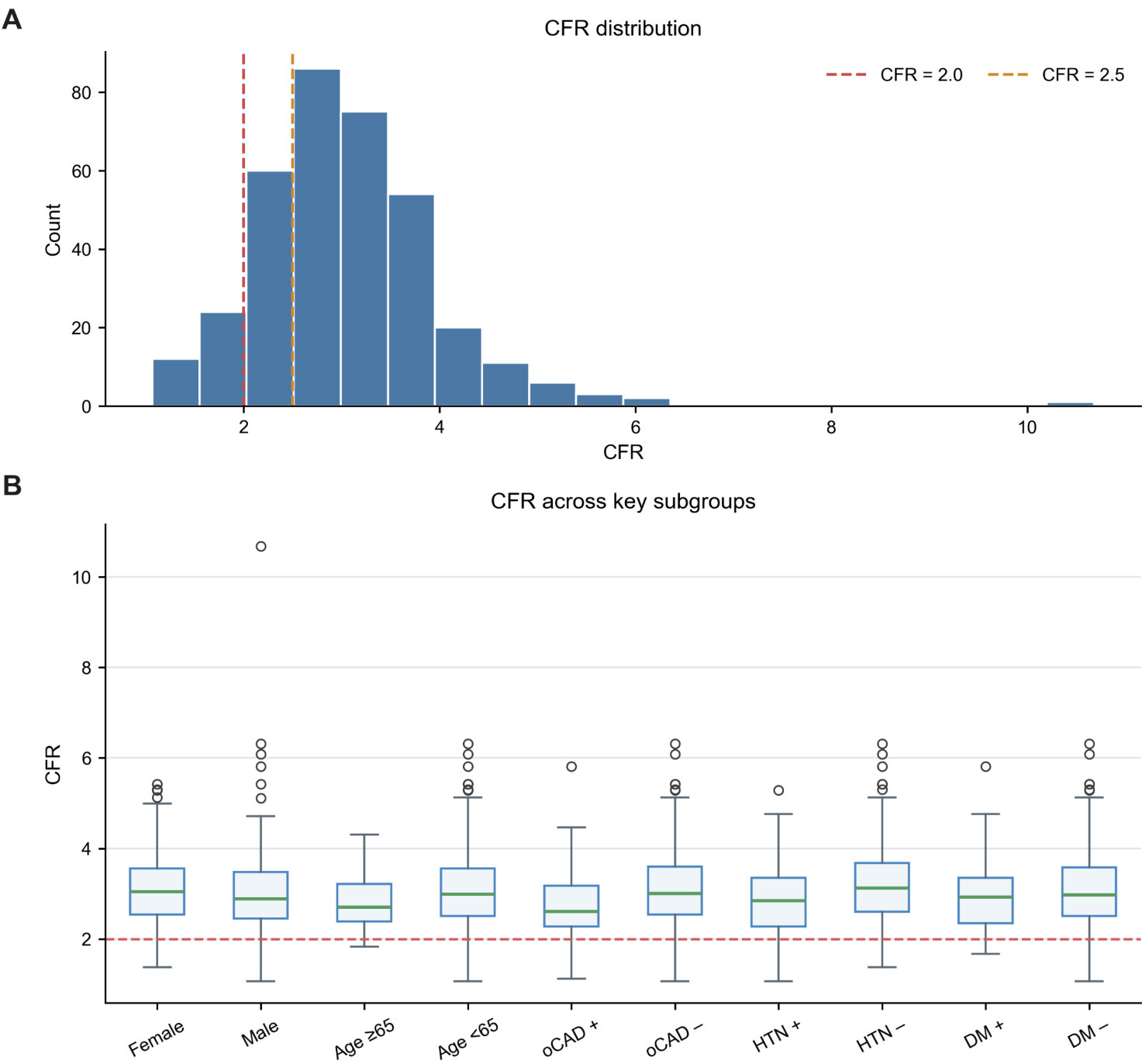


Figure 4

CFR distribution and subgroup overview

Panel A shows the distribution of measured CFR in the full analyzable cohort. Panel B summarizes CFR distributions across prespecified clinical subgroups, including sex, age, oCAD status, hypertension, and diabetes. These descriptive plots provide cohort context and are not intended to imply subgroup-specific model validity.