

Plasma infrared molecular fingerprinting stratifies residual risk in coronary artery disease and reflects clonal haematopoiesis-associated biology

Moritz von Scheidt ^{1,2,*†}, Tarek Eissa ^{1,3,4,*†}, Johannes Krefting^{1,2}, Christian Friess^{1,2}, Marius Schwab^{1,2}, Irina Pugach^{1,2}, Daniel Meyer^{3,4}, Jacqueline Aschauer^{3,4}, Zhifen Chen^{1,2}, Gregor Hoermann⁵, Manja Meggendorfer⁵, Dario Bongiovanni ⁶, Daniel Breucker¹, Karsten Wenger¹, Felix Voll¹, Christian Graesser^{2,7}, Thorsten Kessler^{2,7}, Hendrik B. Sager^{1,2}, Lilith Oldenbuettel¹, Martina Dreßen ^{2,8}, Stefanie A. Doppler^{2,8}, Muamer Keneric⁸, Sabine Seeler^{2,8}, Markus Krane ^{2,8}, Lars Maegdefessel ^{2,9}, Judith S. Hecker^{10,11}, Florian Bassermann^{10,11}, Michaela Sander¹², Stefan Holdenrieder¹², Ingo Hilgendorf ¹³, Florian Leuschner ¹⁴, Shaunak S. Adkar¹⁵, Clint L. Miller¹⁶, Minna U. Kaikkonen¹⁷, Michal Mokry ¹⁸, Gerard Pasterkamp ¹⁸, Pradeep Natarajan ^{19,20,21}, José J. Fuster^{22,23}, Alexander G. Bick²⁴, J. Brett Heimlich²⁴, Michael A. Rieger²⁵, Sebastian Cremer ^{26,27}, Stefanie Dimmeler ^{27,28}, Andreas M. Zeiher^{26,27,28}, Wesley T. Abplanalp^{27,28}, Johan L.M. Björkegren^{29,30,31}, Nicholas J. Leeper ^{15,32}, Birgit Linkohr ^{2,33}, Annette Peters^{2,33,34}, Adnan Kastrati ^{1,2}, Salvatore Cassese ^{1,2}, Wolfgang Koenig ^{1,2}, Frank Fleischmann^{3,4}, Heribert Schunkert^{1,2}, and Mihaela Žigman^{3,4,35,*†}

¹Department of Cardiology, TUM University Hospital German Heart Center, School of Medicine and Health, Technical University of Munich, Lazarettstraße 36, Munich 80636, Germany; ²Deutsches Zentrum für Herz und Kreislaufforschung (DZHK), Partner Site Munich Heart Alliance, Pettenkoferstraße 8a, Munich 80636, Germany; ³Chair of Experimental Physics - Laser Physics, Ludwig-Maximilians-Universität München (LMU), Am Coulombwall 1, Garching 85748, Germany; ⁴Laboratory for Attosecond Physics, Max Planck Institute of Quantum Optics (MPQ), Hans-Kopfermann-Straße 1, Garching 85748, Germany; ⁵MLL Munich Leukemia Laboratory, Munich, Germany; ⁶Department of Internal Medicine I, Cardiology, University Hospital Augsburg, University of Augsburg, Augsburg, Germany; ⁷Department of Internal Medicine III, Cardiology and Angiology, Saarland University Medical Center, Homburg/Saar, Germany; ⁸Department of Cardiovascular Surgery, Institute Insure, TUM University Hospital German Heart Center, School of Medicine and Health, Technical University of Munich, Munich, Germany; ⁹Institute of Molecular Vascular Medicine, TUM Klinikum Rechts der Isar, Technical University Munich, Munich, Germany; ¹⁰Department of Medicine III, TUM Klinikum Rechts der Isar, Munich, Germany; ¹¹TranslaTUM, Center for Translational Cancer Research, Technical University of Munich (TUM), Munich, Germany; ¹²Institute of Laboratory Medicine, TUM Klinikum Deutsches Herzzentrum, Technical University Munich, Munich, Germany; ¹³Department of Cardiology, Angiology and Intensive Care Medicine, Campus Virchow-Klinikum, German Heart Center of the Charité (DHZC), Charité – Universitätsmedizin Berlin, Berlin, Germany; ¹⁴Department of Medicine III, University of Heidelberg, German Centre for Cardiovascular Research (DZHK), Heidelberg, Germany; ¹⁵Division of Vascular Surgery, Department of Surgery, Stanford University School of Medicine, Stanford, CA, USA; ¹⁶Department of Genome Sciences, University of Virginia, Charlottesville, VA, USA; ¹⁷A. I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland; ¹⁸Laboratory of Experimental Cardiology, Department of Cardiology, University Medical Center Utrecht, University Utrecht, Utrecht, The Netherlands; ¹⁹Center for Genomic Medicine and Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA; ²⁰Program in Medical and Population Genetics and the Cardiovascular Disease Initiative, Broad Institute of Harvard and MIT, Cambridge, MA, USA; ²¹Department of Medicine, Harvard Medical School, Boston, MA, USA; ²²Centro Nacional de Investigaciones Cardiovasculares Carlos III, Madrid, Spain; ²³CIBER en Enfermedades Cardiovasculares (CIBER-CV), Madrid, Spain; ²⁴Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; ²⁵Department of Hematology/Oncology, Goethe University Frankfurt, Frankfurt am Main, Germany; ²⁶Department of Cardiology, Goethe University Frankfurt, Frankfurt am Main, Germany; ²⁷German Center for Cardiovascular Research, partner site Frankfurt Rhine-Main, Frankfurt am Main, Germany; ²⁸Institute for Cardiovascular Regeneration, Goethe University Frankfurt am Main, Frankfurt am Main, Germany; ²⁹Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden; ³⁰Department of Genetics and Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ³¹Clinical Gene Networks AB,

* Corresponding author. Emails: moritz.scheidt@tum.de (M.v.S.); Email: tarek.eissa@mpq.mpg.de (T.E.); Email: mihaela.zigman@mpq.mpg.de (M.Ž.)

† These authors contributed equally to the study.

© The Author(s) 2026. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Stockholm, Sweden; ³²Stanford Cardiovascular Institute, Stanford University, Stanford, CA, USA; ³³Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; ³⁴Chair of Epidemiology, Institute for Medical Information Processing, Biometry and Epidemiology, Medical Faculty, LMU, Munich, Germany; and ³⁵Institute of Pathology, Faculty of Medicine, Ludwig-Maximilians-Universität München, Thalkirchner Straße 36, Munich 80337, Germany

Received 28 October 2025; revised 27 March 2026; accepted 13 July 2026

Abstract

Background and Aims

Residual cardiovascular risk remains substantial in patients with coronary artery disease (CAD) despite optimal secondary prevention. Clonal haematopoiesis of indeterminate potential (CHIP) contributes causally to this risk but requires sequencing for detection. Plasma infrared molecular fingerprinting (IMF), a spectroscopy-based assay, was evaluated as a sequencing-free tool for residual risk stratification and CHIP-aligned phenotyping.

Methods

The discovery cohort included 1341 patients with angiographically confirmed CAD (mean age 73 years, 76% male) followed for up to 10.7 years. Plasma IMF spectra, comprehensive clinical data, and targeted sequencing of 13 CHIP driver genes were measured. Penalized Cox models using IMF features were benchmarked against the guideline-recommended SMART2 score. Model performance was assessed by discrimination, reclassification, calibration, and decision curve analyses. External validation was performed in three independent cohorts (PRECAD2, 474 individuals; KORA, 3044 individuals; Lasers4Life, 2123 individuals).

Results

IMF+Age demonstrated superior mortality discrimination over SMART2 (C-index 0.79 vs 0.74, $P < .01$) and significant reclassification improvement. Decision curves showed greater net benefit across clinical thresholds. IMF-derived risk strata showed clear separation of Kaplan–Meier survival curves (log-rank $P < .001$), with mortality increasing from 3% (low-risk) to 27% (high-risk). Higher-risk strata were enriched for CHIP carriers, particularly spliceosome mutations. External validation reproduced expected risk patterns and comorbidity associations.

Conclusions

Plasma IMF provides rapid, low-cost, sequencing-free assessment of residual mortality risk in CAD, outperforming guideline-based scores while reflecting CHIP-associated biology. IMF offers a scalable phenotypic platform for precision secondary prevention and biology-guided trial design.

Structured Graphical Abstract

Key Question

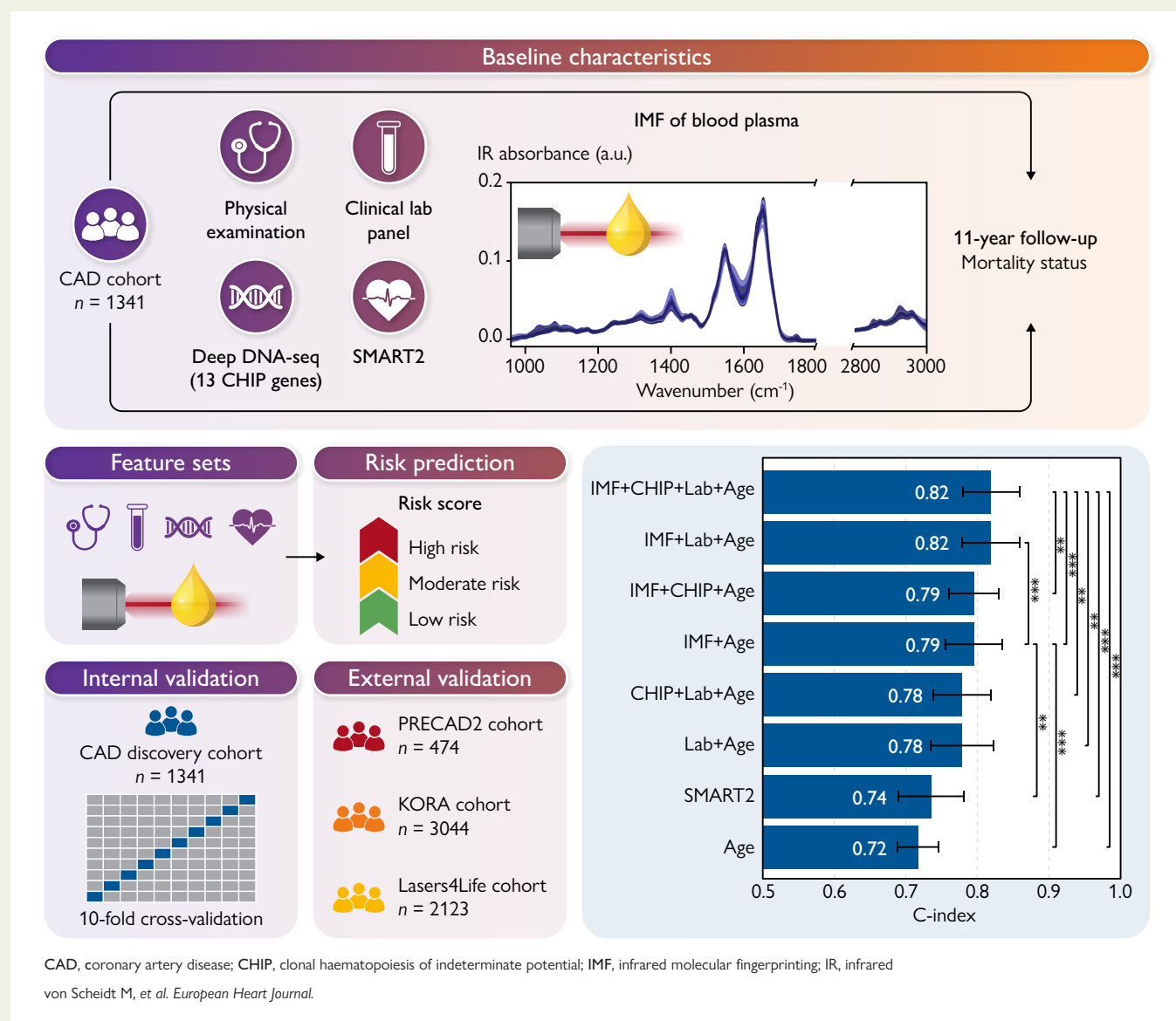
Can plasma infrared molecular fingerprinting (IMF) stratify residual risk in coronary artery disease beyond guideline-recommended scores and capture clonal haematopoiesis of indeterminate potential (CHIP)-associated biological signals?

Key Finding

An IMF+Age model outperformed SMART2 for 11-year mortality prediction with progressive enrichment of CHIP carriers observed among derived risk strata, particularly spliceosome mutations. Findings were validated across three independent cohorts.

Take Home Message

A single, rapid plasma infrared measurement stratifies residual cardiovascular risk without requiring genomic sequencing, enabling scalable identification of high-risk patients for intensified secondary prevention and enrichment of biology-guided clinical trials.



Keywords

Clonal haematopoiesis • Coronary artery disease • Infrared molecular fingerprinting • Spectroscopic profiling • Risk stratification • Precision cardiovascular prevention

Translational perspective

Residual cardiovascular risk persists in patients with coronary artery disease (CAD) despite optimal medical therapy. Detection of clonal haematopoiesis of indeterminate potential (CHIP) identifies individuals at increased risk, but it currently requires next-generation sequencing, which is costly and logistically demanding. Plasma infrared molecular fingerprinting (IMF) provides a rapid, sequencing-free, and scalable approach for mortality risk assessment. In a deeply phenotyped CAD cohort, an IMF+Age model outperformed the guideline-recommended SMART2 score and demonstrated robust calibration, reclassification improvement, and clinical utility across risk thresholds. IMF-derived risk strata aligned with CHIP-associated biology and adverse comorbidity profiles, suggesting capture of systemic inflammatory and haematologic perturbations. Implementation of IMF-based testing could enable efficient identification of high-risk patients for intensified follow-up, optimization of preventive therapy, and enrichment of biology-guided clinical trials in secondary cardiovascular prevention.

Introduction

Residual cardiovascular risk remains substantial in patients with coronary artery disease (CAD) despite intensive guideline-directed secondary prevention. Even with optimal lipid-lowering and antithrombotic therapy, recurrent events and premature mortality continue to occur, underscoring the need for improved tools to identify patients at highest risk and to guide precision prevention strategies.

Clonal haematopoiesis of indeterminate potential (CHIP)—defined by age-related somatic mutations in haematopoietic stem and progenitor cells—has emerged as a causal driver of this residual risk. CHIP promotes systemic inflammation, accelerates atherosclerosis, and confers increased risk of myocardial infarction, stroke, and death.^{1–7} Modulation of interleukin-1 β and interleukin-6 signalling attenuates cardiovascular risk in selected CHIP carriers,^{8,9} supporting its therapeutic relevance. However, CHIP detection currently relies on next-generation sequencing, which is costly, time-consuming, and difficult to scale for routine clinical application. Scalable phenotypic tools that capture CHIP-aligned biology without sequencing could therefore substantially broaden access to precision cardiovascular prevention.

Infrared molecular fingerprinting (IMF) is a spectroscopy-based approach that records vibrational signatures of plasma molecules as they interact with infrared light.^{10,11} By integrating signals from proteins, lipids, and metabolites, IMF generates reproducible composite spectra that reflect systemic physiological states.¹² Prior population-based and oncology studies demonstrated that IMF can detect multimorbid and disease-specific molecular profiles,^{13,14} supporting its potential for minimally invasive health profiling. The method is rapid, low-cost, and requires only microlitre-scale volumes of plasma, allowing application in large-scale clinical and epidemiological settings. A recent study demonstrated the diagnostic potential of infrared spectroscopy in cardiovascular contexts, including cardiac tissue characterization and the optical detection and typing of cardiac amyloidosis from biopsy and fat aspirate samples.¹⁵

We hypothesized that plasma IMF encodes systemic inflammatory, metabolic, and haematologic perturbations enriched in CHIP carriers and linked to residual cardiovascular risk. To test this hypothesis, plasma IMF profiles were analysed in a well-phenotyped CAD cohort with deep CHIP sequencing, long-term mortality follow-up, and comprehensive laboratory assessment. IMF-based models were benchmarked against the guideline-recommended SMART2 algorithm,¹⁶ and mutation-level associations were evaluated with emphasis on spliceosome genes, given their adverse clinical trajectories.⁶ External validation in three

independent IMF-profiled cohorts—a CAD cohort, a population-based cohort, and an oncology cohort^{13,14}—enabled assessment of generalizability and translational feasibility as a sequencing-free biomarker platform for precision secondary prevention.

Methods

Study design, setting, and participants

A prospective cohort study was conducted including 1341 patients (aged 43–94 years, mean age 73 years, 76% male) with angiographically confirmed CAD recruited at TUM University Hospital German Heart Center (Munich, Germany) between 2006 and 2012. All-cause mortality was ascertained via the German national registry, with follow-up of up to 10.7 years. The cohort was previously described,¹⁷ with follow-up now extended from 3 years. Inclusion criteria were stable CAD confirmed by invasive angiography. Key exclusion criteria were acute myocardial infarction at the time of blood sampling, active malignancy, incomplete core clinical data, or follow-up shorter than 1.5 years in survivors. The latter criterion applied to 131 patients (of 1472 initially eligible patients) and was used to ensure adequate outcome ascertainment.

Three independent datasets were used exclusively for external validation of risk models: (i) Personalized Risk Prediction in Coronary Artery Disease 2 (PRECAD2), a prospective study of 474 CAD patients sampled prior to undergoing coronary artery bypass grafting, recruited at TUM University Hospital German Heart Center (Munich, Germany) since 2019 (aged 32–86 years, mean age 69 years, 87% male); (ii) KORA F4, a population-based study of 3044 adults (aged 31–82 years, mean age 56 years, 48% male);^{14,18} and (iii) Lasers4Life, an oncology-focused study of 2123 participants (aged 18–91 years, mean age 64 years, 64% male) including patients with lung, bladder, breast, and prostate cancer as well as controls.^{13,19} All-cause mortality was recorded for all individuals in the PRECAD2 and KORA studies, with follow-up time extending to 7 years in PRECAD2 and 16 years in KORA.

All participants provided written informed consent. The CAD discovery and PRECAD2 studies were approved by the local ethics committee (2013-5943-S-KK, 2018-273-S-KK, 2023-54-S-KK). The KORA F4 study methods were approved by the ethics committee of the Bavarian Chamber of Physicians, Munich (EC No. 06068). The Lasers4Life study participants provided written informed consent for the study under research study protocol #17-141 and under research study protocol #17-182, both of which were approved by the Ethics Committee of the LMU of Munich (German Clinical Trials Register ID DRKS00013217).

Blood collection, processing, and storage

Venous blood was collected into EDTA tubes under standardized conditions, centrifuged within 30 min, plasma aliquoted (50 μ l),

and stored at -80°C until analysis. Samples from the validation cohorts (PRECAD2, KORA, Lasers4Life) underwent comparable processing.

Infrared molecular fingerprinting: acquisition and quality control

Plasma infrared spectra were recorded in their native liquid phase using a Fourier-transform infrared spectrometer (MIRA Analyzer; Clade.io) equipped with a flow-through transmission cuvette (CaF_2 windows; path length $\sim 8\ \mu\text{m}$). Spectra were acquired at $4\ \text{cm}^{-1}$ resolution over $950\text{--}3050\ \text{cm}^{-1}$. Measurements were organized in batches (up to 40 samples) containing pooled human serum as internal laboratory controls to monitor measurement stability. An automated sample delivery system handled sample injection. The mean total acquisition time per sample was approximately 3.5 min, including time for automated flushing in between different samples. Spectral reproducibility was monitored across quality control measurements, with the relative variation being far below 1% across most spectral regions, consistent with prior work,¹¹ confirming high analytical stability.

Spectral preprocessing

Preprocessing of the infrared spectra followed previously established protocols.^{13,14} Spectra were truncated to $1000\text{--}3000\ \text{cm}^{-1}$, the $1800\text{--}2800\ \text{cm}^{-1}$ region was excluded, and each spectrum was L2-normalized. The resulting spectra were treated as feature vectors for all analyses.

CHIP sequencing and variant curation

For the discovery CAD cohort, deep targeted sequencing of 13 canonical CHIP driver genes (DNMT3A, TET2, ASXL1, TP53, SF3B1, SRSF2, U2AF1, PPM1D, JAK2, CALR, CBL, MPL, ZRSR2) was performed at mean coverage exceeding $3500\times$.^{17,20} Variant calling combined Dragen, Pisces, and Pindel pipelines; pathogenicity was adjudicated using COSMIC, ClinVar, and curated in-house databases. CHIP sequencing in the PRECAD2 cohort was performed by deep whole exome sequencing (deepWES) with mean coverage exceeding $500\times$. Variant calling was performed using Mutect2. Variants with variant allele frequency (VAF) $\geq 2\%$ were classified as CHIP.

Clinical variables and reference risk score

Clinical and laboratory variables included creatine kinase, high-sensitivity C-reactive protein (hsCRP), total cholesterol, estimated glomerular filtration rate (eGFR) (Modification of Diet in Renal Disease formula), HDL cholesterol, haemoglobin, haematocrit, creatinine, low-density lipoprotein cholesterol, leukocytes, mean corpuscular volume (MCV), red cell distribution width (RDW), thrombocytes, triglycerides, and troponin T. The SMART2 score was calculated as described by Hageman *et al.*¹⁶ using age, eGFR, hsCRP, smoking, diabetes, non-HDL cholesterol, and sex.

Primary outcome

The primary outcome was all-cause mortality, defined as time from blood draw to death. Survivors were censored at the date of last contact.

Model development and internal validation

Cox proportional-hazards regression with elastic-net regularization was used to derive risk models. In the discovery cohort, predictor sets included (i) age, (ii) SMART2, (iii) IMF features, (iv) clinical laboratory variables, and (v) combined feature sets (e.g. IMF+Lab+CHIP+Age). Prior to model fitting, continuous predictors were

z-score standardized using training split statistics and applied to the corresponding test splits. For the IMF feature set, principal component analysis was applied within each training fold to reduce dimensionality (99.9% cumulative explained variance cut-off, [Supplementary data online, Figure S1](#)), and the resulting components were projected onto the corresponding test splits. Model performance was assessed via a nested 10-fold cross-validation. Within each outer fold, an inner five-fold cross-validation was performed to tune the elastic-net regularization parameters (L1 ratio, alpha). Predicted partial hazard (risk) scores from the held-out samples in each outer split were retained for downstream analyses, yielding a complete set of out-of-sample relative risk scores across all subjects. Regression coefficients and hazard ratios with 95% confidence intervals for the final penalized Cox proportional-hazards models were obtained by refitting each model on the full study cohort after validation to provide stable point estimates. These full-cohort model coefficients are reported for transparency and interpretation and are not identical to the fold-specific models used to generate the cross-validated discrimination estimates. Therefore, caution is warranted when using them for direct individual-level risk calculation without further validation.

Model performance and incremental value analyses

Discrimination was quantified using Harrell's concordance index (C-index). Uncertainty for model discrimination was quantified using non-parametric bootstrap resampling at the subject level. For each model, 2000 bootstrap samples were drawn with replacement from the pooled out-of-sample predictions, and the C-index was recomputed in each sample to obtain 95% confidence intervals (CI). Model comparisons were assessed using paired subject-level bootstrap resampling, and two-sided *P*-values were derived from the empirical distribution of paired C-index differences. Incremental prognostic value over SMART2 was assessed using the time-dependent integrated discrimination improvement (IDI) and the category-free (continuous) net reclassification improvement (NRI) for survival data at 1, 3, 5, and 10 years, computed from out-of-sample Cox predictions.^{21,22} Right-censoring was handled by inverse-probability-of-censoring weighting with the censoring distribution estimated by the Kaplan–Meier method.²³ We report event-NRI, non-event-NRI, and total NRI.²² Uncertainty (95% CIs and two-sided *P*-values) was obtained via paired bootstrap resampling with recomputation of the censoring distribution in each replicate.²³ We additionally evaluated category-based reclassification at 10 years. Absolute risks from each model were mapped to prespecified categories ($<25\%$, $25\text{--}30\%$, $>30\%$) applied identically to the IMF+Age and SMART2 models, and a reclassification table was constructed. The IMF+Age model was selected as the primary risk model for subsequent analyses.

Calibration and decision curve analysis

Model calibration at 1, 3, 5, and 10 years was evaluated by grouping subjects into deciles of predicted risk and comparing the mean predicted risk with the observed event probability estimated via the Kaplan–Meier method. Statistical goodness-of-fit was assessed using the Greenwood–Nam–D'Agostino χ^2 test for survival data, which uses Kaplan–Meier estimates with Greenwood-based variance under censoring. Clinical utility was further examined using decision curve analysis at the same time horizons,²⁴ computing net benefit across threshold probabilities (5%–25%) for the IMF+Age model compared with 'treat all' and 'treat none' strategies, using out-of-sample predictions.

Risk stratification thresholds

Risk groups (low, moderate, high) were defined using maximally selected log-rank statistics ('maxstat'),²⁵ scanning candidate cut-offs between the 10th and 90th percentiles of risk-score distributions in the discovery cohort. The cut-offs yielding maximal log-rank separation were chosen.

Association analyses of comorbidities, biomarkers, and CHIP

Enrichment of categorical variables (CHIP, gene-level mutations, atrial fibrillation, diabetes, hypertension, left ventricular dysfunction, multivessel disease) across IMF risk strata was tested by χ^2 analysis. For continuous markers (e.g. RDW, creatinine, hsCRP, haemoglobin, eGFR), associations were evaluated using Spearman correlation.

Mutation-level analyses

Penalized logistic regression was fit on IMFs to classify carriers vs non-carriers for ASXL1, TET2, and DNMT3A (genes with ≥ 50 carriers) and CHIP at thresholds VAF $\geq 2\%$, $\geq 5\%$, $\geq 10\%$, and $\geq 15\%$. ROC-AUC was estimated by 10-fold stratified cross-validation, reporting the test-set metrics. Age-, sex-, and BMI-matched subsets were used for these classifications.

External validation

The IMF+Age model trained in the CAD cohort was applied without retraining to the independent PRECAD2, KORA, and Lasers4Life cohorts. As a benchmark comparator, the Age-only model was similarly assessed in the external cohorts. To account for demographic differences between the cohorts when calculating the C-index, nearest-neighbour matching (age and sex) was used. Predicted risk-score distributions, derived risk strata, and correlations with comorbidities (hypertension, diabetes, cardiovascular history, and cancer type where available in the external cohorts) were evaluated. In PRECAD2, where CHIP sequencing was performed, associations between the predicted risk strata and CHIP mutations were examined to assess reproducibility with the discovery cohort. Where mortality data were available (PRECAD2, KORA), discrimination of the IMF+Age and Age-only models was quantified using Harrell's C-index with 95% CIs from 2000 subject-level bootstrap resamples, and the two models were compared within each cohort using paired bootstrap resampling, as described above. Calibration and decision curve analyses were performed in the PRECAD2 external validation cohort at 1, 3, and 5 years.

Analysis software

Analyses were conducted in Python (version 3.8.8) using the lifelines package (version 0.27.8) for survival analyses and the scikit-learn package (version 1.3.2) for its implementations of principal component analysis, cross-validation, and logistic regression.

Results

Patient characteristics and CHIP prevalence

A total of 1341 patients with angiographically confirmed CAD (mean age 73 years, 76% male) (Table 1) were included, followed for all-cause mortality for up to 10.7 years. CHIP mutations were detected in 40.6% of participants, consistent with previous CAD cohorts.^{3,4,5,17} DNMT3A (53.1% of carriers), TET2 (35.7%), and ASXL1 (13.4%) were the most frequent, while spliceosome mutations (SRSF2, SF3B1, U2AF1, ZRSR2), although less common, are clinically relevant due to their aggressive clonal

dynamics.^{6,26,27} Across the follow-up period, a total of 143 deaths (10.7%) occurred, corresponding to an incidence rate of 2.4 deaths per 100 person-years.

IMF-based risk prediction outperforms guideline scores

In penalized Cox regression, age alone as a predictor yielded a C-index of 0.72 (95% CI 0.69–0.75) and a SMART2 score of 0.74 (95% CI 0.69–0.78) (Figure 1A). Inclusion of IMF features markedly improved discrimination, with the IMF+Age model achieving a C-index of 0.79 (95% CI 0.75–0.84), significantly higher than SMART2 ($P < .01$) (Figure 1A, Supplementary data online, Table S1). A comprehensive model including IMF, laboratory parameters, age, and CHIP sequencing reached a C-index of 0.82, indicating that IMF captured a major proportion of prognostic information typically obtained from more comprehensive multimodal inputs. Regression coefficients and hazard ratios for all indexed survival models are provided in Supplementary data online, Table S2.

The IMF+Age model demonstrated good overall calibration across 1-, 3-, 5-, and 10-year follow-up periods, with predicted risks closely aligning with observed event rates (see Supplementary data online, Figure S2). Calibration slopes ranged from 1.01 to 1.55 and intercepts from -0.02 to -0.14 , indicating generally accurate risk prediction with a modest pattern of overestimation at lower predicted risks and underestimation at higher predicted risks. Decision curve analysis (see Supplementary data online, Figure S2) showed that the IMF+Age model provided greater net benefit than 'treat all' or 'treat none' strategies across clinically relevant risk thresholds (5%–25%).

Reclassification analyses confirmed significant incremental prognostic value of the IMF+Age model over SMART2 (Figure 1B and C, Supplementary data online, Table S3). The 10-year category-free (continuous) NRI was 0.69 (95% CI 0.35–1.02, $P < .001$), supporting enhanced separation of both survivors and non-survivors by IMF+Age (Figure 1B). This metric is, however, prone to inflation when the new model improves discrimination and is therefore interpreted as supportive. In the category-based reclassification analysis at 10 years, IMF+Age reclassified 570 of 1341 participants (43%) relative to SMART2, comprising 341 (25%) reclassified upward and 229 (17%) downward (see Supplementary data online, Table S3). Observed 10-year Kaplan–Meier event rates in upward-reclassified strata (26%, 42%, 66%) consistently exceeded the overall 10-year event rate of 25%, while downward-reclassified strata (9%, 17%, 19%) fell below it. This indicates that IMF+Age moved participants towards their observed risk and thus reclassified more accurately than SMART2. The 10-year IDI was 0.05 (95% CI 0.03–0.08, $P < .001$, Figure 1C).

Risk stratification and survival differences

Model-derived continuous IMF+Age scores were used to define low-, moderate-, and high-risk strata (Figure 1D). Kaplan–Meier curves demonstrated clear survival probability separation between the three strata (all pairwise log-rank $P < .001$; Figure 1E). Eleven-year mortality was 27% in the high-risk stratum vs 12% and 3% in moderate- and low-risk strata, respectively (see Supplementary data online, Figure S3A). Early mortality within 1 year occurred predominantly in the high-risk group (69%), with no events in the low-risk stratum (see Supplementary data online, Figure S3B and C). These data

Table 1 Baseline characteristics of the coronary artery disease discovery cohort, stratified by clonal haematopoiesis of indeterminate potential status

	All	CHIP+	CHIP–
N	1341	544	797
Age (years)	73.0 ± 9.0	74.2 ± 9.1	72.1 ± 8.9
Male sex	1020 (76.1)	410 (75.4)	610 (76.5)
BMI (kg/m ²)	27.5 ± 4.6	27.5 ± 4.5	27.5 ± 4.6
Diabetes	335 (25.0)	149 (27.4)	186 (23.3)
Hypertension	1162 (86.7)	470 (86.4)	692 (86.8)
Hyperlipidaemia	1194 (89.0)	482 (88.6)	712 (89.3)
Chronic obstructive pulmonary disease	66 (4.9)	35 (6.4)	31 (3.9)
Nicotine intake	520 (38.8)	220 (40.4)	300 (37.6)
History of PCI	868 (64.7)	350 (64.3)	518 (65.0)
History of CABG	194 (14.5)	65 (11.9)	129 (16.2)
History of MI	335 (25.0)	132 (24.3)	203 (25.5)
Family history of CAD	362 (27.0)	153 (28.1)	209 (26.2)
Vessels affected > 1	960 (71.6)	390 (71.7)	570 (71.5)
LV function < 50%	430 (32.1)	191 (35.1)	239 (30.0)
History of CIED	91 (6.8)	38 (7.0)	53 (6.6)
Atrial fibrillation ^a	99 (7.4)	41 (7.5)	58 (7.3)
Atrial fibrillation history ^b	244 (18.2)	100 (18.4)	144 (18.1)
RDW (fl)	44.7 ± 4.4	45.7 ± 5.1	44.0 ± 3.7
Creatinine (mg/dl)	1.09 ± 0.38	1.10 ± 0.31	1.09 ± 0.42
hsCRP (mg/l)	5.0 ± 14.4	5.8 ± 15.7	4.5 ± 13.5
Total cholesterol (mg/dl)	163.1 ± 45.9	157.0 ± 44.2	167.2 ± 46.6
Haematocrit (l/l)	0.40 ± 0.05	0.40 ± 0.06	0.41 ± 0.05
Haemoglobin (g/dl)	13.6 ± 1.7	13.4 ± 1.8	13.8 ± 1.6
GFR (ml/min)	72.4 ± 20.3	70.5 ± 20.0	73.7 ± 20.5
Leukocytes (10 ⁹ /l)	7.5 ± 2.5	7.7 ± 3.0	7.4 ± 2.2
LDL-C (mg/dl)	93.3 ± 40.2	88.7 ± 38.1	96.4 ± 41.3
Thrombocytes (10 ⁹ /l)	212.9 ± 60.3	212.1 ± 61.6	213.6 ± 59.4
Creatine kinase (U/l)	130.8 ± 162.6	134.0 ± 215.8	128.7 ± 113.0
Troponin T (ng/l)	106.0 ± 449.0	125.8 ± 541.0	92.6 ± 373.3
Triglycerides (mg/dl)	129.5 ± 80.1	123.9 ± 74.4	133.4 ± 83.5
HDL-C (mg/dl)	53.5 ± 15.9	53.1 ± 15.8	53.7 ± 16.0
MCV (fl)	89.8 ± 4.9	90.3 ± 5.1	89.4 ± 4.7
Statin therapy	1232 (91.9)	501 (92.1)	731 (91.7)
Anticoagulation therapy	295 (22.0)	121 (22.2)	174 (21.8)
ACE inhibitor/ARB	1128 (84.1)	459 (84.4)	669 (83.9)
Beta-blocker	919 (68.5)	374 (68.8)	545 (68.4)
Calcium antagonist	480 (35.8)	191 (35.1)	289 (36.3)

Continued

Table 1 *Continued*

	All	CHIP+	CHIP–
Platelet inhibitor	1089 (81.2)	432 (79.4)	657 (82.4)
Diuretic therapy ^c	484 (36.1)	201 (36.9)	283 (35.5)
Antidiabetic therapy ^d	350 (26.1)	148 (27.2)	202 (25.3)
ASXL1	75 (5.6)	73 (13.4)	2 (0.3)
CALR	1 (0.1)	1 (0.2)	0 (0.0)
CBL	11 (0.8)	11 (2.0)	0 (0.0)
DNMT3A	349 (26.0)	289 (53.1)	60 (7.5)
JAK2	19 (1.4)	18 (3.3)	1 (0.1)
MPL	4 (0.3)	4 (0.7)	0 (0.0)
PPM1D	49 (3.7)	40 (7.4)	9 (1.1)
SF3B1	27 (2.0)	25 (4.6)	2 (0.3)
SRSF2	26 (1.9)	24 (4.4)	2 (0.3)
TET2	232 (17.3)	194 (35.7)	38 (4.8)
TP53	48 (3.6)	44 (8.1)	4 (0.5)
U2AF1	15 (1.1)	14 (2.6)	1 (0.1)
ZRSR2	13 (1.0)	13 (2.4)	0 (0.0)
Death recorded	143 (10.7)	74 (13.6)	69 (8.7)

Continuous variables are presented as mean $\pm$ standard deviation; binary variables as *n* (%).

Clonal haematopoiesis mutations in the control arm are by definition variant allele frequency < 2%.

ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; BMI, body mass index; CABG, coronary artery bypass grafting; CIED, cardiac implantable electronic device (pacemaker, implantable cardioverter-defibrillator, or cardiac resynchronization therapy device); GFR, glomerular filtration rate (Modification of Diet in Renal Disease formula); HDL-C, high-density lipoprotein cholesterol; hsCRP, high-sensitivity C-reactive protein; LDL-C, low-density lipoprotein cholesterol; LV, left ventricular; MCV, mean corpuscular volume; MI, myocardial infarction; PCI, percutaneous coronary intervention; RDW, red cell distribution width.

^aAtrial fibrillation refers to ECG-documented atrial fibrillation at sampling time.

^bAtrial fibrillation history refers to a prior diagnosis of atrial fibrillation.

^cDiuretic therapy (mineralocorticoid receptor antagonist, loop diuretics, thiazide, or thiazide-like diuretics).

^dAntidiabetic therapy (insulin, biguanides, sodium-glucose cotransporter 2 inhibitors, glucagon-like peptide-1 receptor agonists, dipeptidyl peptidase-4 inhibitors, or sulfonylureas).

highlight the potential of IMF-based stratification to identify vulnerable patients for closer monitoring and intensified secondary prevention.

IMF signatures capture CHIP-associated biology and systemic perturbations

A stepwise enrichment of CHIP carriers was observed from low- to high-risk groups, supporting that IMF reflects systemic molecular phenotypes enriched in individuals with clonal haematopoiesis (*Figure 2A and B*; *Supplementary data online, Table S4*). Among CHIP drivers, splicing factor mutations (SRSF2, SF3B1, U2AF1, ZRSR2) were disproportionately represented in the higher-risk groups, showing stronger relative enrichment than the more common DNMT3A and TET2 variants. Although based on modest numbers, this pattern is consistent with recent studies linking spliceosome mutations to particularly adverse clinical trajectories and accelerated clonal expansion.^{6,26,27} Parallel enrichment of clinical comorbidities—including atrial fibrillation, diabetes, hypertension, reduced left ventricular function, and multivessel disease—was evident across strata, corroborating

the clinical validity of the IMF-based stratification. Laboratory biomarkers further substantiated this stratification (*Figure 2D*, *Supplementary data online, Figure S4*), where higher-risk patients exhibited elevated RDW, creatinine, troponin T, MCV, and hsCRP, along with reduced haemoglobin and glomerular filtration rate (GFR). Together, these associations demonstrate that IMF-derived profiles integrate molecular alterations associated with CHIP and with broader systemic perturbations that collectively drive residual cardiovascular risk.

At the mutation level, IMF signatures demonstrated gene- and clone-size-dependent detectability (*Figure 2D*). ASXL1 mutations showed the highest classification performance (ROC-AUC 0.82), followed by TET2 (0.63) and DNMT3A (0.52), indicating limited mutation-level detection for the most prevalent mutations. Clone-size dependency was observed, with ROC-AUC increasing from 0.58 for CHIP variants with VAF $\geq$ 2% to 0.74 for CHIP variants with VAF $\geq$ 15%. Thus, phenotypic IMF signatures were most pronounced in patients with larger clones or mutations exerting stronger systemic effects.

External validation in the independent PRECAD2 cohort showed similar patterns of enrichment across IMF-derived risk

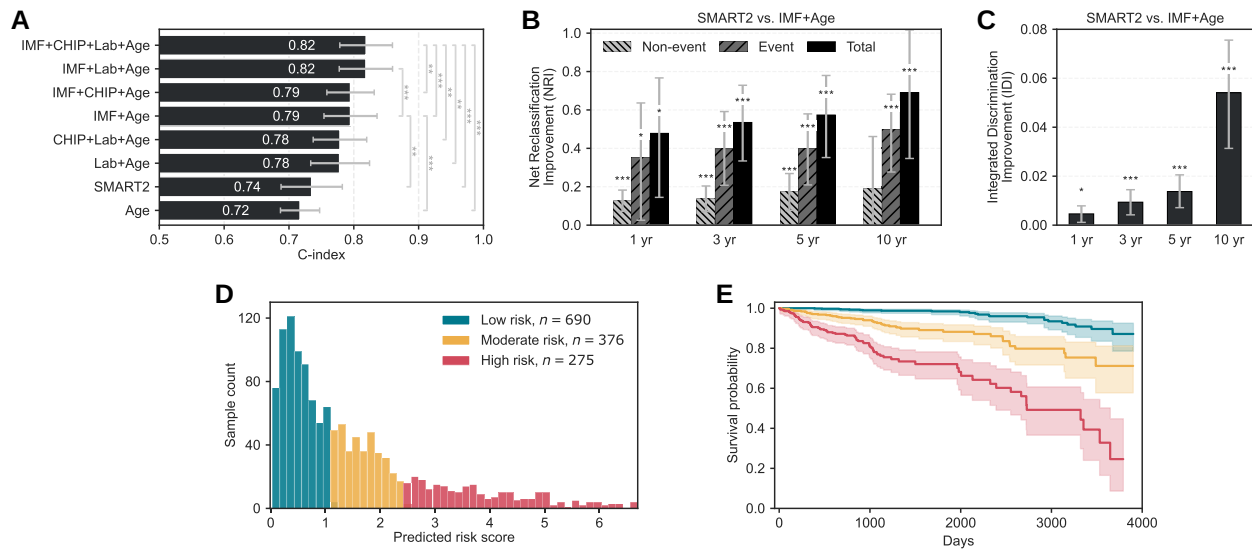


Figure 1 Performance of coronary artery disease risk stratification model. (A) Model discrimination (Harrell's C-index) for different feature sets in the CAD discovery cohort with 95% confidence intervals. Paired statistical comparisons were performed between the best-performing model (IMF+CHIP+Lab+Age) and all other models and between the IMF+Age model and all other models. Significance coded * $P < .05$, ** $P < .01$, *** $P < .001$. (B) Time-dependent continuous net reclassification improvement (NRI) of the IMF+Age model vs the SMART2 model at 1, 3, 5, and 10 years, shown separately for events, non-events, and total. Bar plots indicate point estimates with 95% confidence intervals. (C) Integrated discrimination improvement (IDI, absolute risk difference) of the IMF+Age model vs SMART2 at 1, 3, 5, and 10 years. Bar plots indicate point estimates with 95% confidence intervals. (D) Distribution of continuous IMF+Age risk scores across risk strata. (E) Kaplan-Meier survival curves per IMF+Age-derived risk strata with 95% confidence intervals

strata (Figure 2E and F). CHIP prevalence increased progressively from low- to high-risk groups, with consistent representation of major driver mutations, supporting the reproducibility of the observed associations in an independent CAD population. Although spliceosome mutations were less frequently observed in PRECAD2 due to the smaller cohort size, the detected cases were likewise concentrated in the higher-risk groups.

We further examined IMF spectroscopic patterns of molecular changes for several clinical and genetic variables, with the corresponding effect size curves revealing distinctive molecular fingerprints. Atrial fibrillation showed the greatest qualitative resemblance to the composite risk score, suggesting partial overlap in their underlying molecular perturbations and strengthening the clinical relevance of IMF-derived residual risk stratification (see Supplementary data online, Figure S5A). Some CHIP gene mutations showed qualitative similarities in spectroscopic profiles (see Supplementary data online, Figure S5B), with inferred spectral signatures across different CHIP mutations reflecting underlying biological heterogeneity.

External validation of risk prediction

Application of the IMF+Age model without retraining to the PRECAD2 ($n = 474$; 60 deaths), KORA ($n = 3044$; 472 deaths), and Lasers4Life ($n = 2123$) cohorts reproduced expected risk distributions (Figure 3A). In PRECAD2, 61% of patients were classified as low risk, 24% as moderate risk, and 15% as high risk, demonstrating a similar distribution to the discovery CAD cohort. In KORA, 97% of participants were classified as low risk, consistent with the cohort's predominantly healthy

profile.^{14,18} In Lasers4Life, 22% of individuals were assigned to moderate or high risk, reflecting greater systemic comorbidity and oncologic burden among the participants.^{13,19} Age- and sex-matching of the validation cohorts to the discovery CAD cohort yielded risk-score distributions similar to those observed before matching, with a slight shift towards higher predicted risk (see Supplementary data online, Figure S6). Across the validation cohorts, IMF-derived risk scores correlated with hypertension, diabetes, cardiovascular history, and cancer status, and spectral associations were consistent with those observed in the CAD discovery cohort (see Supplementary data online, Figure S7). The KORA cohort represents a community-based primary-prevention population of younger adults, supporting generalizability of the IMF+Age model beyond secondary prevention.

In PRECAD2, the discovery-derived risk strata showed concordant survival separation, with significantly different outcomes for high vs low and low vs moderate risk (both $P < .001$) and a consistent but non-significant difference for high vs moderate risk ($P = .07$), the latter likely reflecting the smaller number of individuals in the high-risk group (Figure 3B).

In age- and sex-matched subsets of the external cohorts, the IMF+Age model's predictive performance was comparable to that observed in the discovery cohort (C-index 0.74 in PRECAD2 and 0.78 in KORA; Figure 3C). Age alone demonstrated higher discriminatory performance in KORA than in the CAD cohorts, likely reflecting the community-based nature of this population, in which age remains a dominant determinant of mortality risk. In contrast, the CAD cohorts represent secondary-prevention populations with elevated baseline

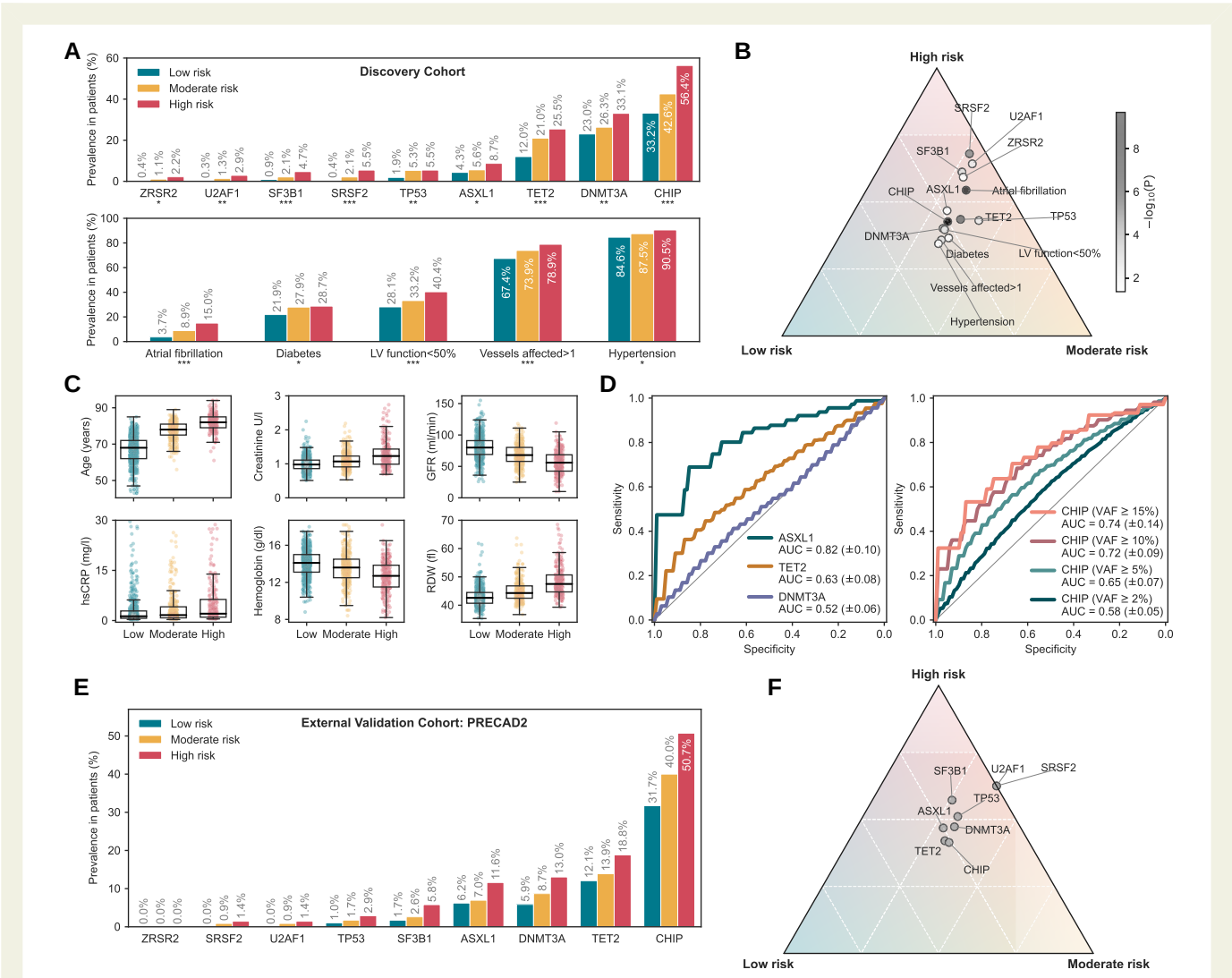


Figure 2 Associations of CHIP and comorbidities across IMF-derived scores. (A) Prevalence of CHIP mutations and comorbidities across IMF+Age model-derived risk strata in the discovery cohort. Significance coded as * $P < .05$, ** $P < .01$, *** $P < .001$. (B) Relative enrichment of CHIP genes and comorbidities across strata, visualized on ternary axes; shading indicates significance (all $P < .05$). (C) Distribution of markers across risk strata shown as boxplots with overlaid individual data points. (D) Area under the curve (AUC) for classifiers trained exclusively on plasma infrared fingerprints to distinguish between carriers and non-carriers across three different mutations and CHIP. Only mutations with at least 50 carriers available in the analysed cohort were considered. CHIP was stratified by variant allele frequency (VAF $< 2\%$ vs VAF $\geq 2\%$, VAF $< 2\%$ vs VAF $\geq 5\%$, VAF $< 2\%$ vs VAF $\geq 10\%$, and VAF $< 2\%$ vs VAF $\geq 15\%$). Matched cohorts for age, sex, and body mass index are detailed in [Supplementary data online, Table S5](#). (E) External validation of associations between CHIP mutations and risk strata in the PRECAD2 cohort. The IMF+Age model trained in the discovery cohort was applied to PRECAD2 cohort data. The bar plot shows the prevalence of CHIP mutations across risk strata. (F) Relative enrichment of CHIP genes across risk strata in the PRECAD2 external validation cohort

cardiovascular risk across age groups, thereby attenuating the prognostic contribution of chronological age alone. This was most evident in PRECAD2, where the Age-only model yielded a lower C-index (0.61). PRECAD2 represented the smallest external cohort and had the shortest effective follow-up with the fewest observed mortality events, resulting in greater statistical uncertainty. In addition, PRECAD2 exclusively included patients undergoing CABG, representing a clinically more homogeneous high-risk population with a narrower age distribution than the discovery cohort and KORA. Despite the weaker performance of age alone, the IMF+Age model maintained robust

discrimination in PRECAD2 (C-index 0.74), supporting the incremental prognostic value contributed by IMF-derived molecular information, particularly in clinically homogeneous CAD populations.

Summary of findings

Across analyses, plasma IMF provided a robust, sequencing-free predictor of long-term mortality in CAD. The IMF+Age model outperformed the guideline-based SMART2 score across discrimination, calibration, reclassification (IDI/NRI), and decision

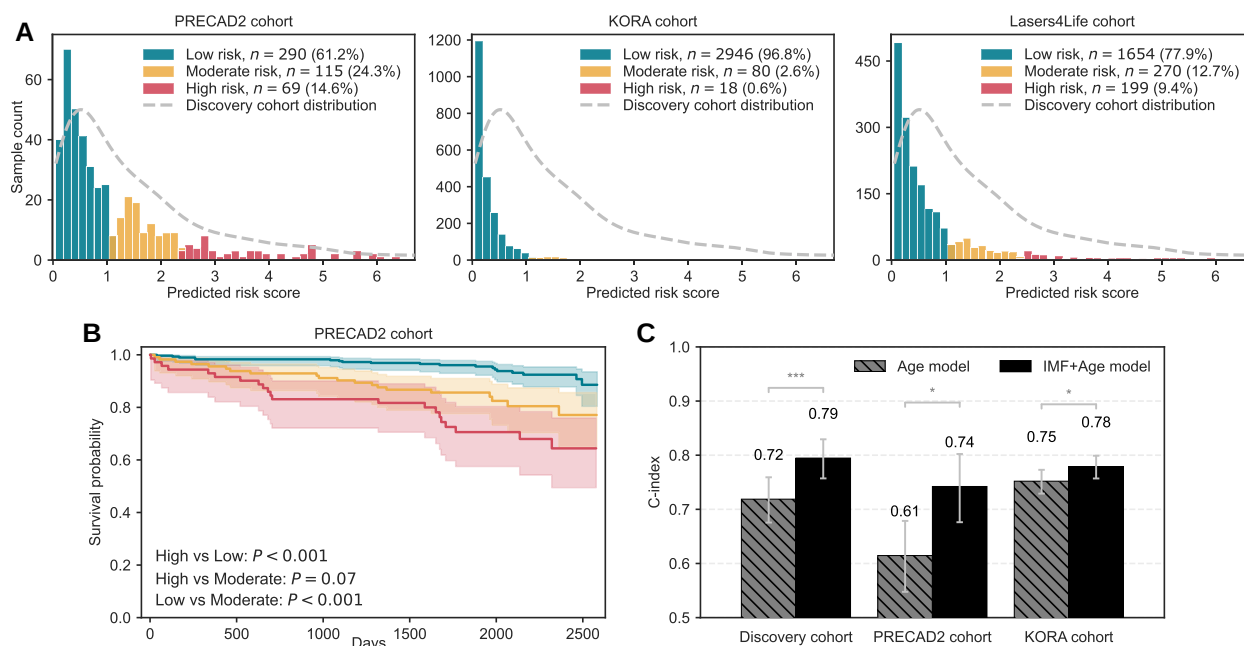


Figure 3 External validation across three independent cohorts. (A) Distributions of predicted risk scores generated by the IMF+Age model trained in the CAD discovery cohort and applied without retraining to the PRECAD2 CAD patient cohort, the population-based KORA cohort, and the oncology cohort Lasers4Life. Risk strata were defined using thresholds derived from the discovery cohort. The dashed line shows the discovery cohort risk score distribution on an arbitrary y-axis for reference. (B) Kaplan–Meier survival curves per IMF+Age-derived risk strata with 95% confidence intervals in the PRECAD2 external validation cohort; P -values derived from pairwise log-rank tests. (C) Model discrimination (Harrell’s C-index) of the IMF+Age model compared with an Age-only model in the discovery cohort and in the external PRECAD2 and KORA cohorts. Models trained in the discovery cohort were applied without retraining to age- and sex-matched subsets of PRECAD2 and KORA matched to the discovery cohort. Lasers4Life was not included as mortality data were unavailable. Error bars indicate 95% confidence intervals determined by subject-level bootstrap resampling of test-set predictions

curve utility. IMF-derived risk strata integrated CHIP status, comorbidity burden, and systemic biomarker profiles and generalized to three external cohorts, supporting its reproducibility and translational feasibility as a scalable platform for precision secondary prevention.

Discussion

In a large, well-characterized CAD cohort with long-term follow-up, plasma IMF enabled stratification of residual mortality risk without the need for sequencing. IMF profiles also reflected biology associated with CHIP. An easy-to-employ IMF+Age model outperformed the guideline-recommended SMART2 score for 10-year mortality prediction and maintained superior calibration, reclassification, and decision curve net benefit. IMF-derived risk strata were associated with significantly different survival trajectories, aligned with CHIP status—particularly spliceosome mutations—and generalized across three independent cohorts ([Structured Graphical Abstract](#)). IMF is not intended as a platform to detect CHIP mutations *per se*. Instead, IMF presents a scalable phenotypic readout integrating haematologic, inflammatory, and metabolic determinants of residual cardiovascular risk, including CHIP-associated systemic consequences.

Advances beyond prior work

Previous studies have shown that CHIP contributes causally to incident cardiovascular events and mortality, with spliceosome mutations conferring particularly adverse trajectories.^{3–6,26,27} In parallel, optical spectroscopy has demonstrated diagnostic potential across medical applications.^{10–13,15} The present study bridges these domains by demonstrating that a single, rapid plasma measurement captures CHIP-related systemic alterations and predicts long-term mortality beyond a validated guideline score. IMF achieved higher discrimination, calibration, and reclassification and retained performance across external populations. The finding that IMF+Age alone approximated the performance of multimodal models including laboratory and sequencing data underscores its capacity to capture clinically actionable biological variation.

Plasma IMF records vibrational signatures embedded across circulating proteins, lipids, and metabolites and therefore cannot directly measure genomic-level perturbations. Accordingly, genomic alterations can only be reflected indirectly through downstream effects on the circulating molecular milieu. Phenotypic profiling with proteomics, lipidomics, and metabolomics provides complementary molecular resolution. In coronary heart disease, protein-, lipid-, and metabolite-based risk models have reported C-statistics ranging from 0.63 to 0.78 (follow-up

periods of 1–10 years),^{28–30} placing both the IMF+Age C-index of 0.79 and the comprehensive IMF+CHIP+Lab+Age model C-index of 0.82 in a comparable or favourable range. However, direct cross-study comparison is limited by differences in endpoints, cohorts, and analytic frameworks. A practical advantage of IMF is that a single, rapid, low-cost measurement inherently integrates information across proteins, lipids, and metabolites, without requiring separate high-throughput assays for each molecular layer. Studies assessing IMF alongside other molecular profiling platforms on the same samples are required to robustly benchmark their relative and combined prognostic value for residual cardiovascular risk stratification.

Biological plausibility

Several complementary mechanisms support IMF as a phenotypic surrogate of CHIP-related pathophysiology. First, chronic activation of IL-1 β /IL-6 signalling in CHIP reshapes the acute-phase proteome and lipoprotein landscape, altering mid-infrared absorption profiles. Second, increased oxidative stress and metabolic reprogramming modify protein secondary structures and glycation states captured by IMF.^{10,11,31} Experimental work has shown that CHIP-associated IL-1 β and IL-6 signalling accelerates atherosclerosis and systemic inflammation, supporting the mechanistic plausibility of this pathway.^{32,33} Third, spliceosome mutations induce cytopenic and dysmetabolic states that affect erythropoiesis, iron turnover, and plasma lipid composition, with these alterations encoded in the infrared spectrum.^{26,27} Finally, systemic comorbidities such as renal dysfunction, anaemia, and inflammation—frequent in CHIP carriers—further shape IMF patterns. The observed gradients in hsCRP, RDW, creatinine, and troponin across IMF risk strata are consistent with this integrative, multi-pathway capture of cardiovascular vulnerability.

Clinical relevance and implementation

IMF requires microlitre-scale plasma, minimal preparation, and yields results within minutes on an automated platform, enabling same-visit risk assessment. In clinical workflows, IMF+Age could identify high-risk patients for intensified surveillance, optimization of guideline-directed therapy, or enrolment in biology-guided trials. Such trials may include anti-inflammatory strategies targeting the IL-1 β /IL-6 axis, where CHIP carriers appear to derive particular benefit.^{8,9} IMF could also serve as a front-line screening tool to prioritize patients for confirmatory sequencing or mechanistic investigations, thereby improving efficiency and reducing cost barriers. Because IMF relies on standardized, untargeted signal acquisition rather than individual analytes, it is inherently suited for longitudinal monitoring and for use in laboratories where sequencing infrastructure is unavailable. In current secondary-prevention frameworks, 10-year mortality risks above 10%–15% are considered clinically actionable thresholds for intensified management. IMF+Age demonstrated the greatest net benefit within this range.

Limitations and future directions

Several limitations should be acknowledged. CHIP sequencing was available only in the discovery cohort and one CAD external validation cohort, precluding mutation-specific validation in the

population-based and oncological cohorts. Future studies should determine whether plasma IMF provides information that is complementary to, rather than redundant with, proteomic, glycomic, metabolomic, lipidomic, and other molecular profiling approaches in CAD. Without integrated multi-omic analyses, the biological interpretation of individual spectral features remains indirect. The study population consisted predominantly of older individuals with established CAD, warranting validation in younger, multi-ethnic, and primary-prevention cohorts. Inter-platform and inter-site reproducibility require formal evaluation to facilitate clinical translation. Exclusion of survivors with follow-up shorter than 1.5 years may have introduced selection bias by removing early losses to follow-up. Risk stratification thresholds were selected *post hoc* on pooled out-of-sample risk scores. The reported low/moderate/high group separation in the discovery cohort is therefore subject to selection optimism, although a nested cross-validation sensitivity analysis indicated the magnitude of this optimism was small, with cut-off points largely unchanged. Finally, whether IMF-guided risk stratification can improve clinical outcomes should be tested prospectively in pragmatic or stepped-wedge designs.

Conclusions

Plasma IMF provides a rapid, scalable, and sequencing-free readout of residual cardiovascular risk that aligns with CHIP-associated biology and systemic comorbidities. By improving discrimination, calibration, and clinical utility over a guideline-based score, IMF represents a feasible bridge between omics-based insight and routine clinical care, enabling precision secondary prevention and efficient enrichment of biology-guided cardiovascular trials.

Acknowledgements

We thank all study participants and acknowledge the Cardiovascular Biobank of the TUM University Hospital German Heart Center, Technical University Munich, as well as the clinical, laboratory, and research personnel involved in the conduct of the study. We thank the members of the Broadband Infrared Diagnostics team at LMU Physics for their support and feedback.

Supplementary data

Supplementary data are available at [European Heart Journal](#) online.

Declarations

Disclosure of Interest

P.N. reports research grants from Allelica, Amgen, Apple, Boston Scientific, Cleerly, Genentech/Roche, Ionis, Novartis, and Silence Therapeutics; consulting or speaker fees from AIRNA, Allelica, Apple, AstraZeneca, Bain Capital, Blackstone Life Sciences, Bristol Myers Squibb, Creative Education Concepts, CRISPR Therapeutics, Eli Lilly & Co, Esperion Therapeutics, Foresite Capital, Foresite Labs, Genentech/Roche, GV,

HeartFlow, Incyte, Magnet Biomedicine, Merck, Novartis, Novo Nordisk, TenSixteen Bio, and Tourmaline Bio; equity in Bolt, Candela, Mercury, MyOme, Parameter Health, Preciseli, and TenSixteen Bio; royalties from Recora; and spousal employment at Vertex Pharmaceuticals. All other authors declare no conflicts of interest.

Data Availability

Pseudonymized individual-level data from the CAD cohort are available through the German CHIP Registry e.V. (<https://chip-registry.com/>) upon reasonable application and approval. Access to KORA and Lasers4Life data is available through their data access committees under controlled conditions (KORA, <https://helmholtz-muenchen.de/en/epi/kora>; Lasers4Life, <https://www.lasers4life.de>). Custom Python analysis code is available upon request.

Funding

This work was supported by LMU Munich Centre for Advanced Laser Applications (CALA), the non-profit German CHIP Registry (Deutsches CHIP Register e.V., www.chip-register.de), the DZG Innovation fund (Inter-Organ Metabolomics, 'MetaboCHIP', DZHK-81X2600534), the German Centre for Cardiovascular Research (DZHK; DZHK-81X2200145, DZHK-81Z0600501, DZHK-81X2600520, DZHK-81X3600506), the German Heart Foundation (Deutsche Herzstiftung e.V.), the CORONA Foundation (S199/10085/2021), the Fondation Leducq (PlaQOmics, AteroGen consortia), the Bavarian State Ministry of Health and Care (DigiMed Bayern, DMB-1805-0001), the European Union's Horizon Europe European Innovation Council programme (Grant Agreement No. 101115381), the European Union and the Innovative Health Initiative Joint Undertaking (Grant agreement No. 101194785), the German Research Foundation (DFG-95/1713-1), the Bavarian State Ministry of Science and the Arts Joint Research Center Deutsches Herzzentrum München and Munich School of Robotics and Machine Learning), the German Federal Ministry of Education and Research (ERA-CVD framework), the Swedish Research Council, the Swedish Heart Lung Foundation, the National Institutes of Health/NHLBI (R01HL164577, R01HL148167, R01HL148239, R01HL166428, R01HL168174), the American Heart Association (19TPA34910021), and the Accelerating Medicines Partnership Program for Common Metabolic Diseases, managed by the Foundation for the National Institutes of Health.

Ethical Approval

All participants provided written informed consent. The study was approved by the local ethics committee (2013-5943-S-KK, 2018-273-S-KK, 2023-54-S-KK).

Pre-registered Clinical Trial Number

None supplied.

References

- Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 2014;371:2488–98. <https://doi.org/10.1056/nejmoa1408617>
- Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* 2014;371:2477–87. <https://doi.org/10.1056/nejmoa1409405>
- Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N Engl J Med* 2017;377:111–21. <https://doi.org/10.1056/nejmoa1701719>
- Bick AG, Pirruccello JP, Griffin GK, Gupta N, Gabriel S, Saleheen D, et al. Genetic interleukin 6 signaling deficiency attenuates cardiovascular risk in clonal hematopoiesis. *Circulation* 2020;141:124–31. <https://doi.org/10.1161/circulationaha.119.044362>
- Diez-Diez M, Ramos-Nebbe BL, de la Barrera J, Silla-Castro JC, Quintas A, Vázquez E, et al. Unidirectional association of clonal hematopoiesis with atherosclerosis development. *Nat Med* 2024;30:2857–66. <https://doi.org/10.1038/s41591-024-03213-1>
- Gumuser ED, Schuermans A, Cho SMJ, Sporn ZA, Uddin MM, Paruchuri K, et al. Clonal hematopoiesis of indeterminate potential predicts adverse outcomes in patients with atherosclerotic cardiovascular disease. *J Am Coll Cardiol* 2023;81:1996–2009. <https://doi.org/10.1016/j.jacc.2023.03.401>
- Rauch PJ, Gopakumar J, Silver AJ, Nachun D, Ahmad H, McConkey M, et al. Loss-of-function mutations in Dnmt3a and Tet2 lead to accelerated atherosclerosis and concordant macrophage phenotypes. *Nat Cardiovasc Res* 2023;2:805–18. <https://doi.org/10.1038/s44161-023-00326-7>
- Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N Engl J Med* 2017;377:1119–31. <https://doi.org/10.1056/NEJMoa1707914>
- Marston NA, Pirruccello JP, Melloni GEM, Kamanu F, Bonaca MP, Giugliano RP, et al. Clonal hematopoiesis, cardiovascular events and treatment benefit in 63,700 individuals from five TIMI randomized trials. *Nat Med* 2024;30:2641–7. <https://doi.org/10.1038/s41591-024-03188-z>
- Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, et al. Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* 2014;9:1771–91. <https://doi.org/10.1038/nprot.2014.110>
- Huber M, Kepesidis KV, Voronina L, Božić M, Trubetskov M, Harbeck N, et al. Stability of person-specific blood-based infrared molecular fingerprints opens up prospects for health monitoring. *Nat Commun* 2021;12:1511. <https://doi.org/10.1038/s41467-021-21668-5>
- Eissa T, Voronina L, Huber M, Fleischmann F, Žigman M. The perils of molecular interpretations from vibrational spectra of complex samples. *Angew. Chem. Int. Ed* 2024;63:e202411596. <https://doi.org/10.1002/anie.202411596>
- Huber M, Kepesidis KV, Voronina L, Fleischmann F, Fill E, Hermann J, et al. Infrared molecular fingerprinting of blood-based liquid biopsies for the detection of cancer. *eLife* 2021;10:e68758. <https://doi.org/10.7554/elife.68758>
- Eissa T, Leonardo C, Kepesidis KV, Fleischmann F, Linkohr B, Meyer D, et al. Plasma infrared fingerprinting with machine learning enables single-measurement multi-phenotype health screening. *Cell Rep Med* 2024;5:101625. <https://doi.org/10.1016/j.xcrm.2024.101625>
- Mukherjee SS, Maleszewski JJ, Luthringer D, Confer MP, Mittal A, Dasari S, et al. Optical spectroscopic detection and typing of cardiac amyloidosis. *Circulation* 2024;150:1299–301. <https://doi.org/10.1161/circulationaha.124.069554>
- Hageman SHJ, McKay AJ, Ueda P, Gunn LH, Jernberg T, Hagström E, et al. Estimation of recurrent atherosclerotic cardiovascular event risk in patients with established cardiovascular disease: the updated SMART2 algorithm. *Eur Heart J* 2022;43:1715–27. <https://doi.org/10.1093/eurheartj/ehac056>
- von Scheidt M, Adkar SS, Krefting J, Hoermann G, Meggendorfer M, Bauer S, et al. Clonal haematopoiesis of indeterminate potential and mortality in coronary artery disease. *Eur Heart J* 2026;47:453–69. <https://doi.org/10.1093/eurheartj/ehaf602>
- Linkohr B, Heier M, Gieger C, Thorand B, Grallert H, Holle R, et al. Cohort profile: Cooperative Health Research in the Region of Augsburg (KORA) 1984–2024. *Int J Epidemiol* 2025;54:dyaf187. <https://doi.org/10.1093/ije/dyaf187>
- Kepesidis KV, Jacob P, Schweinberger W, Huber M, Feiler N, Fleischmann F, et al. Electric-field molecular fingerprinting to probe cancer. *ACS Cent Sci* 2025;11:560–73. <https://doi.org/10.1021/acscentsci.4c02164>
- Cremer S, von Scheidt M, Kirschbaum K, Tombor L, Mas-Peiro S, Abplanalp W, et al. Prognostic significance of somatic mutations in myeloid cells of men with chronic heart failure – interaction between loss of Y chromosome and clonal haematopoiesis. *Eur J Heart Fail* 2025;27:3219–34. <https://doi.org/10.1002/ehf.3778>
- Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–72. <https://doi.org/10.1002/sim.2929>
- Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11–21. <https://doi.org/10.1002/sim.4085>

23. Uno H, Tian L, Cai T, Kohane IS, Wei LJ. A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. *Stat Med* 2013;**32**:2430–42. <https://doi.org/10.1002/sim.5647>
24. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;**26**:565–74. <https://doi.org/10.1177/0272989X06295361>
25. Hothorn T, Lausen B. On the exact distribution of maximally selected rank statistics. *Comput Stat Data Anal* 2003;**43**:121–37. [https://doi.org/10.1016/s0167-9473\(02\)00225-6](https://doi.org/10.1016/s0167-9473(02)00225-6)
26. Niroula A, Sekar A, Murakami MA, Trinder M, Agrawal M, Wong WJ, *et al.* Distinction of lymphoid and myeloid clonal hematopoiesis. *Nat Med* 2021;**27**:1921–7. <https://doi.org/10.1038/s41591-021-01521-4>
27. Robertson NA, Latorre-Crespo E, Terradas-Terradas M, Lemos-Portela J, Purcell AC, Livesey BJ, *et al.* Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. *Nat Med* 2022;**28**:1439–46. <https://doi.org/10.1038/s41591-022-01883-3>
28. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, *et al.* Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. *JAMA* 2016;**315**:2532–41. <https://doi.org/10.1001/jama.2016.5951>
29. Hilvo M, Meikle PJ, Pedersen ER, Tell GS, Dhar I, Brenner H, *et al.* Development and validation of a ceramide- and phospholipid-based cardiovascular risk estimation score for coronary artery disease patients. *Eur Heart J* 2020;**41**:371–80. <https://doi.org/10.1093/eurheartj/ehz387>
30. Lv J, Pan C, Cai Y, Han X, Wang C, Ma J, *et al.* Plasma metabolomics reveals the shared and distinct metabolic disturbances associated with cardiovascular events in coronary artery disease. *Nat Commun* 2024;**15**:5729. <https://doi.org/10.1038/s41467-024-50125-2>
31. Voronina L, Fleischmann F, Šimunović J, Ludwig C, Novokmet M, Žigman M. Probing blood plasma protein glycosylation with infrared spectroscopy. *Anal Chem* 2024;**96**:2830–9. <https://doi.org/10.1021/acs.analchem.3c03589>
32. Sano S, Oshima K, Wang Y, MacLauchlan S, Katanasaka Y, Sano M, *et al.* Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1 β /NLRP3 inflammasome. *J Am Coll Cardiol* 2018;**71**:875–86. <https://doi.org/10.1016/j.jacc.2017.12.037>
33. Fuster JJ, MacLauchlan S, Zuriaga MA, Polackal MN, Ostriker AC, Chakraborty R, *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* 2017;**355**:842–7. <https://doi.org/10.1126/science.aag1381>