# Identification of Growth Differentiation Factor-15 as An Early Predictive Biomarker for Metabolic Dysfunction-Associated Steatohepatitis: A Nested Case-control Study of UK Biobank Proteomic Data

Hao Wang[1,2], Xiaoqian Xu[1,2], Yameng Sun[1,2], Hong You[1,2], Jidong Jia[1,2], You-Wen He[3], Yuanyuan Kong[1,2]

[1] National Clinical Research Center for Digestive Disease, State Key Lab of Digestive Health, Beijing Friendship Hospital, Capital Medical University, Beijing, China.

[2] Department of Clinical Epidemiology and Evidence-Based Medicine, Beijing Clinical Research Institute, Beijing, China.

[3] Department of Integrative Immunobiology, Duke University School of Medicine, Durham, NC 27710, USA

**Correspondence to:**

Prof. Yuanyuan Kong, PhD

National Clinical Research Center for Digestive Diseases, State Key Lab of Digestive Health, Beijing Friendship Hospital, Capital Medical University; Clinical Epidemiology and EBM Unit, Beijing Clinical Research Institute, Beijing, China.

E-mail: kongyy@ccmu.edu.cn.

## Abstract

**Background/Aims:** This study aims to determine the predictive capability for metabolic dysfunction-associated steatohepatitis (MASH) long before its diagnosis by using six previously identified diagnostic biomarkers for metabolic dysfunction-associated steatotic liver disease (MASLD) with proteomic data from the UK Biobank.

**Methods:** A nested case-control study comprising of a MASH group and three age- and sex-matched controls groups (metabolic dysfunction-associated steatosis, viral hepatitis, and normal liver controls) were conducted. Olink proteomics, anthropometric and biochemical data at baseline levels were obtained from the UK Biobank. The baseline levels of CDCP1, FABP4, FGF21, GDF15, IL-6 and THBS2 were analyzed prospectively to determine their predictive accuracy for subsequent diagnosis with a mean lag time of over 10 years.

**Results:** At baseline, GDF15 demonstrated the best performance for predicting MASH occurrence at 5 and 10 years later, with an AUC of 0.90 at 5 years and 0.86 at 10 years. A predictive model based on four biomarkers (GDF15, FGF21, IL-6, and THBS2) showed AUCs of 0.88 at both 5 and 10 years. Furthermore, a protein-clinical model that included these four circulating protein biomarkers along with three clinical factors (BMI, ALT and TC) yielded AUCs of 0.92 at 5 years and 0.89 at 10 years.

**Conclusions:** GDF15 at baseline levels outperformed other individual circulating protein biomarkers for the early prediction of MASH. Our data suggest that GDF15 and the GDF15-based model may be used as easy-to-implement tools to identify patients with high risk of developing MASH at a mean lag time of over 10 years.

**Keywords:** Metabolic dysfunction-associated steatohepatitis, Metabolic dysfunction-associated steatotic liver disease, Growth differentiation factor 15, Predictive biomarkers, Proteomics

**Highlights:**

- The predictive capabilities of six previously identified circulating proteins (CDCP1, FABP4, FGF21, GDF15, IL-6, and THBS2) for metabolic dysfunction-associated steatohepatitis (MASH) long before its diagnosis were analyzed prospectively in a nested case-control study by using proteomic data from the UK Biobank.
- Growth differentiation factor 15 (GDF15) exhibited superior predictive capability for MASH development with significant accuracy even 5 to 10 years prior to onset in individuals with normal liver function.
- GDF15 and the GDF15-based model may be used as easy-to-implement tools to identify patients with high risk of developing MASH at a mean lag time of over 10 years.

## Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD), formerly known as non-alcoholic fatty liver disease (NAFLD), affects about 38% of the global population and is the most common chronic liver disease worldwide.[1] MASLD encompasses a spectrum of disease severity, including hepatic steatosis with or without mild inflammation (metabolic dysfunction-associated steatosis, MAS) and a necroinflammatory subtype (metabolic dysfunction-associated steatohepatitis, MASH). MASH, the active form of MASLD, is characterized by the histological inflammation and faster fibrosis progression.[2] The global prevalence of MASH is 5.27%,[3] and it has become the second or third leading cause of end-stage liver disease and hepatocellular carcinoma (HCC).[4,5] A large proportion of MASH patients remain undiagnosed until the disease has progressed to more advanced and life-threatening stages due to its largely asymptomatic nature in the early stage.[6] Thus, early prediction of patients at high risk of MASH is an important medical need and will play a significant role in the timely intervention and personalized treatment from the perspective of preventive medicine.

To date, histological assessment of liver biopsy remains the gold standard for diagnosing MASH. However, the limitations of biopsy, including its invasiveness, associated complications, high cost, time-consuming nature and intra-observer variability, make it less suitable for MASH monitoring in clinical practice, and unsuitable for screening in the population-based health surveys.[7] Imaging based diagnostic tools, such as ultrasonography and magnetic resonance imaging, are

primarily used for detecting MASH, but these tests cannot reliably differentiate patients with MASH from patients with MAS and lack predictive capability for MASH disease development.[8] Therefore, developing innovative and non-invasive biomarkers for predicting MASH holds a promising prospect for clinical application.

During the last decade, six non-invasive protein biomarkers have been identified as diagnostic biomarkers to differentiate patients with MASH, MASLD, and advanced fibrosis.[9,10] These biomarkers include CUB domain-containing protein 1 (CDCP1), fatty acid binding protein 4 (FABP4), fibroblast growth factor 21 (FGF21), growth differentiation factor 15 (GDF15), interleukin-6 (IL-6), and thrombospondin 2 (THBS2). CDCP1 is a transmembrane glycoprotein involved in the tyrosine phosphorylation-dependent regulation of cellular events related to tumor invasion and metastasis.[11] Recent omics-based studies have identified serum soluble CDCP1 as a robust serological biomarker for risk stratification of patients with MASH.[12] FABP4 is a fatty acid binding protein that participates in fatty acid uptake, transport, and metabolism.[13] A previous study found that FABP4 level could be a potential biomarker for MASLD diagnosis.[14] FGF21 is a stress-inducible hormone that has important roles in regulating energy balance and glucose and lipid homeostasis. FGF21 is a potential pharmaceutical target for obesity-related metabolic complications, including hyperglycemia, dyslipidemia, and MASH.[15] GDF15 is a member of the transforming growth factor β (TGF-β) superfamily whose expression increases in response to cellular stress and diseases.[16] GDF-15 levels rise with the progression of MASLD severity.[17] IL-6 is a cytokine that functions in inflammation and regulation of metabolic

homeostasis.[18] Increasing evidence indicates that IL-6 is involved in metabolic diseases, including MASLD and MASH.[19,20] THBS2 is a homotrimeric glycoprotein mediating cell-to-cell and cell-to-matrix interactions, and a key determinant of fibrogenesis in MASLD.[21] These identified diagnostic biomarkers have shown potential in distinguishing between different stages and severities of liver disease, highlighting their significance in the early prediction and management of MASH.

Based on the identification of the above six circulating diagnostic biomarkers for MASLD and MASH, we hypothesize that these soluble non-invasive biomarkers may serve as predictive biomarkers for MASH development long before disease diagnosis. To test this hypothesis, we analyzed the proteomic data from the prospective UK Biobank study which recruits human subjects, establishes their multi-omics at baseline levels, and then follows disease development in the entire cohort.

## Materials and Methods

### Study design

To validate proteomic biomarkers within the UK Biobank, a nested case-control study was conducted. This approach involves selecting cases and controls from the population of an existing cohort, offering several advantages over traditional case-control studies. The study was approved by the North West Multi-Centre Research Ethics Committee, and detailed information about its methodology and protocols can be found on the official UK Biobank website. The current study's protocol is available online and conducted under the application number 129053.

**Definition of cases and multiple controls**

Based on the International Classification of Diseases 10th revision (ICD-10), four distinct groups were identified from hospital medical records: MASH, MAS, viral hepatitis (VH), and normal liver (NL) groups. Below is how each group was defined and selected for comparison:

**MASH group:**

Participants diagnosed with MASH were defined as the case group. Consistent with the Expert Panel Consensus statement, MASH cases were defined as ICD-10 K75.8 and were excluded if they had: (1) a prior diagnosis of MASH/MASLD before blood sample collection; (2) a diagnosis of MASH/MASLD within six months from recruitment; or (3) a diagnosis of other liver diseases (**Table S1**).

**Control groups:**

Three control groups were established for comprehensive comparison: MAS, VH, and NL groups. MAS group was defined as ICD-10 K76.0 and were excluded if they had the same exclusion criteria as MASH group. VH group was defined as ICD-10 B15.0, B16.0, B17.0, B18.0, B19.0, and were excluded if they had: (1) a prior diagnosis of VH before blood sample collection; or (2) a diagnosis of other liver diseases. Participants without any medical records of liver diseases were defined as NL controls.

**Matching Protocol:**

MASH cases and non-MASH controls were matched on age (within the same year) and sex in a 1:2 ratio to minimize confounding factors. Due to sample size constraints,

VH controls were matched with MASH cases in a 1:1 ratio. Non-MASH controls were randomly selected from their respective groups to ensure unbiased comparison. This matching and selection strategy aimed to elucidate the unique proteomic profiles associated with each condition (MASH, MAS, VH, and NL), facilitating a comprehensive analysis of biomarkers and their potential diagnostic utility in differentiating these liver disease states within the UK Biobank cohort. Comorbidities of each group at baseline were presented in Table 1.

**Anthropometric and Biochemical Measurements**

Age at baseline was determined from the participants' dates of birth and the dates of their baseline assessments. Anthropometric measurements, including height, weight, and body mass index (BMI) were collected. Routine biochemical tests were obtained from the UK Biobank (detailed in Supplementary Methods). The FIB-4 index was calculated as (age (year) * AST [U/L]) / ((PLT [$10^9$/L]) * (ALT [U/L])1/2). This data, along with the proteomic data, was used to analyze and validate the six candidate protein biomarkers (CDCP1, FABP4, FGF21, GDF15, IL-6, and THBS2) for their predictive capability in identifying individuals at high risk of developing MASH. Additional information on UK Biobank study population and Olink proteomics is included in Supplemental Methods.

**Statistical analysis**

Continuous variables were presented as means ± standard deviations (SD) or

medians (interquartile ranges) based on their distributions. Differences among case and control groups were tested using one-way ANOVA or the Kruskal-Wallis test. For comparisons between two groups, the Student's t-test or Mann-Whitney *U*-test were used. Categorical variables were presented as frequencies and percentages and analysed using the Chi-square test. To adjust for multiple comparisons, the Bonferroni correction was applied, with the adjusted alpha level calculated as $\alpha adjust=\alpha_0/\kappa$ [22]. Spearman's correlation was employed to estimate the associations between candidate proteomic biomarkers and clinical factors.

Multivariable Cox regression analyses were performed to identify independent proteomic biomarkers for MASH, adjusting for three different models. Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated to estimate the risk effects of candidate proteomic biomarkers for MASH. To further improve the predictive accuracy of potential protein biomarkers for MASH, Cox regression models were performed to construct predictive models based on the candidate protein biomarkers and MASH-related clinical risk factors. The time-dependent receiver operating characteristic (ROC) curves, which measure the performance of candidate biomarkers given the disease status of individuals at certain time points, were used to evaluate the predictive performance of the models. The area under the curve (AUC) was calculated to assess the predictive performance of the models. The performance of the predictive models was evaluated using 10-fold cross-validation on the training set. Statistical analyses were performed using R, and two-tailed $P < 0.05$ was considered statistically significant.

## Results

### Baseline characteristics of the participants

A total of 53,014 participants with proteome data were obtained from UK Biobank. According to their ICD-10 codes, 82 MASH cases, 831 MAS cases, 114 VH cases and 51,987 NL controls were identified. Among the 82 MASH cases, 4 patients with MASH at baseline were excluded from this study. Among the 831 MAS cases, 68 patients with MAS at baseline, 9 patients with other concurrent liver diseases, and 7 patients diagnosed with MAS within six months were excluded from this study. Based on the matching criteria, a total of 78 MASH cases and three groups of age- and sex-matched controls, including 156 MAS cases, 78 VH cases, and 156 NL controls, were included. This selection process ensures that the study groups are well-matched and minimizes potential confounding factors, thereby improving the reliability of the comparative analyses (**Fig. 1A**).

The mean times from baseline to diagnosis for MASH, MAS and VH cases are 12.93, 11.59, and 12.63 years, respectively (**Fig. 1B**). The demographic, anthropometric, and biochemical characteristics of participants are described in **Table 1**. The mean ages of MASH, MAS, VH, and NL control groups were approximately 59 years.

Baseline liver function tests (ALT, AST, FIB-4, and ALP) were significantly elevated in the MASH group, but nevertheless remained within the normal ranges when compared to the NL control group at baseline levels (all $P < \alpha_{adjust}$, 0.01) (**Table 1**). No

significant differences were observed between the two groups for ALB, PLT and TBiL levels (all $P > 0.01$). Significant differences in carbohydrate and lipid metabolism markers, including BMI, CRP, TC, TG, and HDL, were noted between the MASH and NL groups (all $P < 0.01$). Kidney function markers, such as UA and urea, showed no significant differences between the MASH and NL groups (all $P > 0.01$). No significant differences were observed in terms of the reported metabolism or digestive system comorbidities, including diabetes, hypertension, digestive system cancers, gastrointestinal tract and biliary tract diseases between the MASH and NL groups (all $P > 0.01$). These findings suggest that the comorbidities at baseline were largely comparable (**Table 1**).

Significant differences were identified in BMI, ALT, AST, FIB-4, FPG, TC, and TG, but not in other parameters in the MASH group when compared to the VH group (**Table 1**). Taken together, these baseline characteristics suggest liver functional changes in patients with MASH long before their corresponding disease diagnosis.

Importantly, no significant differences were observed between the MASH and MAS groups in liver function and carbohydrate and lipid metabolism, except for the FIB-4 and GGT baseline levels. These baseline characteristics suggest that liver function and carbohydrate and lipid metabolism in MASH and MAS cases were similar at the baseline levels, indicating that these parameters are not sufficient to differentiate MASH cases from patients with MAS (**Table 1**).

**Expression levels of six proteomic biomarkers in MASH at baseline**

We compared the levels of six proteins among the four groups and found that the expression levels of circulating CDCP1, FABP4, FGF21, GDF15, IL-6, and THBS2 at baseline were highest in the MASH group (all $P < 0.01$), followed by a gradually decreasing trend in MAS, VH, and NL control groups (**Table 2 and Fig. 2**).

Significantly higher levels of all six protein biomarkers were identified in the MASH group compared to the VH and NL control groups (All $P < 0.01$). Except for FABP4 and FGF21, other four biomarkers were significantly higher in the MASH group compared to the MAS group at baseline levels (All $P < 0.01$). All six protein biomarkers were significantly higher in the MAS group compared to NL controls at baseline levels (All $P < 0.01$). There were no significant differences for CDCP1, FABP4, FGF21, and IL-6, between VH group and NL controls (All $P > 0.01$, **Fig. 2**).

These findings highlight the potential of these six protein biomarkers, particularly GDF15, CDCP1, IL-6, and THBS2, to serve as early indicators of MASH. The elevated levels of these biomarkers in MASH patients long before the clinical diagnosis suggest their potential utility in the early detection and risk stratification of this disease.

**Association among baseline levels of the six circulating biomarkers and clinical characteristics**

To determine the relationships among the six biomarkers, Spearman's correlation coefficients were calculated and visualized in a correlation plot (**Fig. 3A**). The baseline levels of all biomarkers exhibited positive correlations with each other ($P < 0.05$). Notably, GDF15 and CDCP1 demonstrated the strongest positive correlation ($r = 0.59$,

$P < 0.05$), suggesting a high degree of association between these two potential biomarkers.

We also calculated the correlation coefficients between protein biomarkers and liver function, kidney function, and metabolic tests. The six protein biomarkers displayed positive correlations with ALT, AST, ALP, GGT and FIB-4 (**Fig. 3**). CDCP1, FABP4, GDF15, IL-6, and THBS2 showed negative correlations with ALB (**Fig. 3A**). Furthermore, the six biomarkers at baseline levels positively correlated with carbohydrate and lipid metabolism markers, including BMI, FPG, CRP, and TG, and negatively correlated with TC and HDL (**Fig. 3B**), highlighting their potential relevance to MASH pathology. Taken together, these correlation analyses underscore the interconnected nature of these biomarkers and their association with liver function and metabolic health.

**Risk effects of the six circulating protein biomarkers for MASH**

To determine the risk effects of these circulating biomarkers for MASH development prior to diagnosis, we calculated the HRs and 95% CIs using Cox regression. Three models were employed to adjust the clinical factors to remove potential confounding bias. Model 1 adjusted for one metabolic marker and two liver function markers, including BMI, ALT, and AST. Model 2 adjusted for BMI, ALT, AST, ALB, ALP, GGT, and PLT. Model 3 adjusted for factor in model 2 and TC, TG, HDL, and LDL.

Among the six biomarkers, GDF15 at baseline levels showed the highest HR for

discriminating MASH from NL controls. After adjustment using model 1, high GDF15 baseline level was associated with a roughly 2.5-fold increased risk of MASH (HR = 2.49, $P < 0.001$). Further adjustments using model 2 revealed that high GDF15 baseline level was associated with a 2.9-fold increased risk of MASH (HR = 2.89, $P < 0.001$). After additional adjustments using model 3, the HR for GDF15 decreased slightly but remained highly significant (HR = 2.338, $P < 0.001$, **Fig. 4A**). The elevated levels of the other five protein biomarkers were also associated with an increased the risk of MASH when compared to NL controls, though their risk effects were not as strong as GDF15 (**Fig. 4A**).

Only two biomarkers, FGF21 and THBS2, were associated with increased risk of MAS development (**Fig. 4B**). None of the six circulating biomarkers at baseline levels were associated with increased risk of VH (**Fig. 4C**). In conclusion, even after adjusting for clinical factors, GDF15 remained the most predictive biomarker for MASH, indicating its potential utility in early prediction and intervention strategies.

**Predictive performance of the six biomarkers at baseline for subsequent MASH development**

GDF15 at baseline level demonstrated the best performance for predicting MASH occurrence against NL controls. Specifically, the AUC for GDF15 at 5 years was 0.90, and at 10 years, it was 0.86 (**Fig. 5**). Further analysis involved constructing predictive models based on the six protein biomarkers using Cox regression models. The best model was selected using the stepwise (both directions) Akaike's Information Criterion

method. This process identified four key biomarkers (GDF15, FGF21, IL-6, and THBS2) for the final predictive model. The predictive performance of this model showed an AUC of 0.88 at both 5 years and 10 years (**Fig. 5**).

To enhance predictive accuracy further, a protein-clinical model was developed by incorporating clinical factors at baseline levels along with the protein biomarkers. This model included four key biomarkers (GDF15, FGF21, IL-6, and THBS2) and three clinical factors (BMI, ALT and TC), resulting in an AUC of 0.92 at 5 years and 0.89 at 10 years (**Fig. 5**). The predictive performances of the routinely used clinical parameters, including ALT, BMI, TC and FIB-4, yielded AUCs lower than 0.9. As the commonly used non-invasive index to evaluate liver fibrosis, FIB-4 showed lower AUCs than GDF15 and two combinatory models (**Fig. 5**).

Importantly, GDF15 also exhibited the best performance for predicting MASH occurrence against MAS at 5 and 10 years, with AUCs of 0.85 and 0.78, respectively. The protein model for MASH occurrence showed AUCs of 0.82 at 5 years and 0.77 at 10 years, while the protein-clinical model showed AUCs of 0.86 at 5 years and 0.80 at 10 years (**Fig. 6**). To assess GDF15 and two combined algorithms performed in general populations, time-dependent ROC curves were performed in unmatched populations (78 MASH and 50827 NL controls). GDF15 at baseline level also showed best performance for predicting MASH against NL controls (**Fig. S1**). Furthermore, GDF15 at baseline level demonstrated strong predictive accuracy for distinguishing MASH from VH, with AUCs of 0.84 at 5 and 0.77 at 10 years (**Fig. S2**).

The predictive performance of these protein biomarkers for MAS were also

evaluated individually, but none yielded AUCs higher than 0.7 for predicting MAS occurrence against NL controls at 5 or 10 years (**Fig. S3**). Predictive models for MAS, based on the protein biomarkers and clinical factors, also showed relatively poor performance. The protein model for MAS had AUCs of 0.619 at 5 years and 0.719 at 10 years, while the protein-clinical model had AUCs of 0.672 at 5 years and 0.734 at 10 years (**Fig. S3**). Additionally, the performance of the six individual biomarkers and developed predictive models for VH was inferior to their performance for predicting MASH and MAS, highlighting the excellent specificity of these biomarkers for MASH prediction (**Fig. S4**).

In conclusion, GDF15 at baseline demonstrated superior predictive accuracy for MASH development, outperforming the other five biomarkers as well as clinically used index marker FIB-4. The combination of protein biomarkers and clinical factors further enhanced predictive performance, with GDF15 playing a pivotal role.

## Discussion

In this study, we investigated the predictive accuracy of six previously identified diagnostic biomarkers including CDCP1, FABP4, FGF21, GDF15, IL-6, and THBS2 for predicting MASH development using proteomic data from a prospective cohort in the UK Biobank. Our investigation identified GDF15 at baseline levels as the most effective predictive biomarker for stratifying patients at high risk of developing MASH compared to those who later developed MAS, VH, or individuals with normal liver function. Furthermore, we developed two models with high sensitivities and

specificities for predicting MASH based on these circulating protein biomarkers. Our findings suggest that proteomics biomarkers hold promising clinical applications in the prediction and prevention of MASH.

GDF15, also known as macrophage inhibitory cytokine-1, plays a crucial role in regulating TGF-β signalling.[23] It is closely involved in infection, fibrosis, and apoptosis pathways in response to tissue damage or disease progression.[16] In the context of MASH, liver fibrosis is a significant hallmark, where hepatic stellate cells (HSCs) play a pivotal role. TGF-β and related signals can accelerate HSC activation, promoting their proliferation, migration, contraction, and secretion of extracellular matrix proteins, thereby contributing to liver fibrogenesis.[24] Like other TGF-β superfamily cytokines, GDF15 plays an important role in pro-fibrotic effect in liver fibrogenesis.[25] A multicentre transcriptomic study demonstrated that hepatic GDF15 expression was positively associated with MASLD severity, and GDF15 expression was significantly higher in patients with advanced liver fibrosis.[26] Taken together, GDF15 may promote liver fibrogenesis in MASH progression, and the circulating level of GDF15 could be elevated long before disease diagnosis of MASH.

Clinically, the detected range of GDF15 in blood of large cohorts of old adults (>70 year old) and patients with chronic obstructive pulmonary disease by immunoassay method is about 1.1-1.4 μg/L,[27-29] and by proximity extension assay method is about 3.0-5.0 normalized protein expression (NPX, **Table 2**).[30] Our study found that the mean levels of GDF15 in MASH, MAS, and NL control groups is 1.63, 1.18 and 0.02 NPX, respectively, well below the above reported 3.0-5.0 NPX value. Furthermore, a

previous study found that as diagnostic biomarkers, the mean levels of GDF15 by immunoassay in MASH patients, MAS patients and non-MASLD controls is 1.15 μg/L 0.97 μg/L, and 0.62 μg/L, respectively.[17] The mean levels of GDF15 in none/mild fibrosis (F0-1), moderate (F2), and severe fibrosis (F3-4) groups is 0.78 μg/L, 0.93 μg/L, and 1.50 μg/L, respectively.[17] A previous case-control study detected serum GDF15 using immunoassay and found that GDF15 is the most important single biomarker for the diagnosis of MASLD patients with moderate-to-severe fibrosis (F2-4 fibrosis).[31] The serum level of GDF15 is significantly higher in MASLD patients with advanced liver fibrosis, and it can distinguish F2-4 MASLD patients with a sensitivity of 0.56, a specificity of 0.87, and an AUC of 0.75.[31] The diagnostic cut-off value based on Youden index is 1.19 μg/L.[31] Our study found that GDF15 at baseline levels at 1.63 NPX is significantly increased in MASH patients, and it can predict MASH occurrence from different control groups with high sensitivity and specificity. In contrast, none of the elevated clinical factors at baseline levels can predict MASH development with high specificity. Comparing to serving as a diagnostic biomarker for MASH and MASLD, our findings suggest that GDF15 at baseline level could be used as a predictive biomarker for MASH long before its diagnosis. In future study, it will be important to measure the GDF15 at baseline using immunoassay to quantitate the concentration of GDF15.

Currently, many clinical diagnostic models and blood-based biomarkers have been developed as alternatives for the non-invasive diagnosis of patients with MASH.[9] Several these non-invasive diagnostic biomarkers may also predict adverse clinical

outcomes, such as liver-related mortality, cardiovascular events, HCC, or liver transplantation.[32,33] However, no highly sensitive and specific tests are available to predict MASH long before the disease diagnosis. The lack of MASH predictive biomarkers may be due to the slow progression from simple hepatic steatosis to MASH, and only a part of patients with simple hepatic steatosis can progress to MASH. Therefore, developing predictive biomarkers for MASH needs a long-term follow-up cohort with large sample size. The UK Biobank, with multi-omics data and long-term follow-up of large size participants, presents an invaluable resource for identifying accurate predictive biomarkers for MASH.

Our study, utilizing proteome profiling, uniquely identifies GDF15 as a robust circulating protein biomarker for early MASH prediction, demonstrating its predictive efficacy with a mean lag time exceeding 10 years. While acknowledging limitations such as sample size and the need for further multi-centered validation, our findings suggest GDF15 and GDF15-based models as promising tools for predictive diagnosis and population risk stratification in MASH.

In conclusion, GDF15 emerges as a promising biomarker for MASH prediction, surpassing other studied protein biomarkers. Its incorporation into easy-to-implement predictive models could potentially reduce the need for invasive procedures like liver biopsies, benefiting both patients and healthcare systems alike. Continued research and validation efforts are essential to fully harness the predictive potential of GDF15 in clinical practice.

## Acknowledgments

The authors acknowledge the participants and their families who contributed to make this study possible. This work was partially supported by the research grants from National Key Research and Development Program, China (2023YFC2306902, 2023YFC2306900), High-level Public Health Technical Talents of the Beijing Municipal Health Commission, China (XUEKEGUGAN-010-018), Beijing Municipal Administration of Hospitals Incubating Program (PX2024003), Research Grant of Capital Medical University (PYZ21051, PYZ23073).

## Conflicts of interest

The authors confirm that there are no conflicts of interest.

## Abbreviations

AUC, area under the curve; CDCP1, CUB domain-containing protein 1; FABP4, fatty acid binding protein 4; FIB-4, fibrosis-4 index; FGF21, fibroblast growth factor 21; GDF15, growth differentiation factor 15; HCC, hepatocellular carcinoma; HR, hazard ratio; HSCs, hepatic stellate cells; ICD-10, International Classification of Diseases 10th revision; NL, normal liver controls; NPX, normalized protein expression; ROC, receiver operating characteristic; TGF-β, transforming growth factor β; THBS2, thrombospondin 2; VH: viral hepatitis.

## Authors' contribution

Conceptualization: HY, JJ, YWH, and YK; Methodology: HW and YK; Software: HW and YS; Validation: XX; Formal analysis: HW; Investigation: XX and YS; Data curation: HW; Writing-original draft preparation: HW; Writing-review and editing: YWH and YK; Visualization: HW; Supervision: JJ, and YK; Funding acquisition: HY and YK. All authors have read and agreed to the published version of the manuscript.

## Supplementary material

Supplementary material is available at Clinical and Molecular Hepatology website (http://www.e-cmh.org).

## References


1. Wong VW, Ekstedt M, Wong GL, Hagstrom H. Changing epidemiology, global trends and implications for outcomes of NAFLD. J Hepatol 2023;79:842-852.
2. Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clin Gastroenterol Hepatol 2015;13:643-654.
3. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology 2023;77:1335-1347.
4. Huang DQ, El-Serag HB, Loomba R. Global epidemiology of NAFLD-related HCC: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2021;18:223-238.
5. Goldberg D, Ditah IC, Saeian K, Lalehzari M, Aronsohn A, Gorospe EC, et al. Changes in the Prevalence of Hepatitis C Virus Infection, Nonalcoholic Steatohepatitis, and Alcoholic Liver Disease Among Patients With Cirrhosis or Liver Failure on the Waitlist for Liver Transplantation. Gastroenterology 2017;152:1090-1099.
6. Balp MM, Krieger N, Przybysz R, Way N, Cai J, Zappe D, et al. The burden of non-alcoholic steatohepatitis (NASH) among patients from Europe: A real-world patient-reported outcomes study. JHEP Rep 2019;1:154-161.
7. Rinella ME, Sanyal AJ. Management of NAFLD: a stage-based approach. Nat Rev Gastroenterol Hepatol 2016;13:196-205.
8. Ajmera V, Loomba R. Imaging biomarkers of NAFLD, NASH, and fibrosis. Mol Metab 2021;50:101167.
9. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. J Hepatol 2018;68:305-315.
10. Wong VW, Adams LA, de Ledinghen V, Wong GL, Sookoian S. Noninvasive

biomarkers in NAFLD and NASH - current progress and future promise. Nat Rev Gastroenterol Hepatol 2018;15:461-478.

11. Lim SA, Zhou J, Martinko AJ, Wang YH, Filippova EV, Steri V, et al. Targeting a proteolytic neoepitope on CUB domain containing protein 1 (CDCP1) for RAS-driven cancers. J Clin Invest 2022;132:e154604.
12. Jia X, Song E, Liu Y, Chen J, Wan P, Hu Y, et al. Identification and multicentric validation of soluble CDCP1 as a robust serological biomarker for risk stratification of NASH in obese Chinese. Cell Rep Med 2023;4:101257.
13. Hotamisligil GS, Bernlohr DA. Metabolic functions of FABPs--mechanisms and therapeutic implications. Nat Rev Endocrinol 2015;11:592-605.
14. Coilly A, Desterke C, Guettier C, Samuel D, Chiappini F. FABP4 and MMP9 levels identified as predictive factors for poor prognosis in patients with nonalcoholic fatty liver using data mining approaches and gene expression analysis. Sci Rep 2019;9:19785.
15. Geng L, Lam KSL, Xu A. The therapeutic potential of FGF21 in metabolic diseases: from bench to clinic. Nat Rev Endocrinol 2020;16:654-667.
16. Wang D, Day EA, Townsend LK, Djordjevic D, Jorgensen SB, Steinberg GR. GDF15: emerging biology and therapeutic applications for obesity and cardiometabolic disease. Nat Rev Endocrinol 2021;17:592-607.
17. Valenzuela-Vallejo L, Chrysafi P, Kouvari M, Guatibonza-Garcia V, Mylonakis SC, Katsarou A, et al. Circulating hormones in biopsy-proven steatotic liver disease and steatohepatitis: A Multicenter Observational Study. Metabolism 2023;148:155694.
18. Giraldez MD, Carneros D, Garbers C, Rose-John S, Bustos M. New insights into IL-6 family cytokines in metabolism, hepatology and gastroenterology. Nat Rev Gastroenterol Hepatol 2021;18:787-803.
19. Cobbina E, Akhlaghi F. Non-alcoholic fatty liver disease (NAFLD) - pathogenesis, classification, and effect on drug metabolizing enzymes and transporters. Drug Metab Rev 2017;49:197-211.
20. Park J, Zhao Y, Zhang F, Zhang S, Kwong AC, Zhang Y, et al. IL-6/STAT3 axis

dictates the PNPLA3-mediated susceptibility to non-alcoholic fatty liver disease. J Hepatol 2023;78:45-56.

21. Kimura T, Iwadare T, Wakabayashi SI, Kuldeep S, Nakajima T, Yamazaki T, et al. Thrombospondin 2 is a key determinant of fibrogenesis in non-alcoholic fatty liver disease. Liver Int 2024;44:483-496.
22. Sedgwick P. Multiple hypothesis testing and Bonferroni's correction. BMJ 2014;349:g6284.
23. Wischhusen J, Melero I, Fridman WH. Growth/Differentiation Factor-15 (GDF-15): From Biomarker to Novel Targetable Immune Checkpoint. Front Immunol 2020;11:951.
24. Hammerich L, Tacke F. Hepatic inflammatory responses in liver fibrosis. Nat Rev Gastroenterol Hepatol 2023;20:633-646.
25. Qi P, Ma MZ, Kuai JH. Identification of growth differentiation factor 15 as a pro-fibrotic factor in mouse liver fibrosis progression. Int J Exp Pathol 2021;102:148-156.
26. Govaere O, Cockell S, Tiniakos D, Queen R, Younes R, Vacca M, et al. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. Sci Transl Med 2020;12:eaba4448.
27. Daniels LB, Clopton P, Laughlin GA, Maisel AS, Barrett-Connor E. Growth-differentiation factor-15 is a robust, independent predictor of 11-year mortality risk in community-dwelling older adults: the Rancho Bernardo Study. Circulation 2011;123:2101-2110.
28. Eggers KM, Kempf T, Wallentin L, Wollert KC, Lind L. Change in growth differentiation factor 15 concentrations over time independently predicts mortality in community-dwelling elderly individuals. Clin Chem 2013;59:1091-1098.
29. Martinez CH, Freeman CM, Nelson JD, Murray S, Wang X, Budoff MJ, et al. GDF-15 plasma levels in chronic obstructive pulmonary disease are associated with subclinical coronary artery disease. Respir Res 2017;18:42.

30. Zhong W, Edfors F, Gummesson A, Bergstrom G, Fagerberg L, Uhlen M. Next generation plasma proteome profiling to monitor health and disease. Nat Commun 2021;12:2493.
31. Bilson J, Scorletti E, Bindels LB, Afolabi PR, Targher G, Calder PC, et al. Growth differentiation factor-15 and the association between type 2 diabetes and liver fibrosis in NAFLD. Nutr Diabetes 2021;11:32.
32. Armandi A, Sanavia T, Younes R, Caviglia GP, Rosso C, Govaere O, et al. Serum ferritin levels can predict long-term outcomes in patients with metabolic dysfunction-associated steatotic liver disease. Gut 2024;73:825-834.
33. Mozes FE, Lee JA, Vali Y, Alzoubi O, Staufer K, Trauner M, et al. Performance of non-invasive tests and histology for the prediction of clinical outcomes in patients with non-alcoholic fatty liver disease: an individual participant data meta-analysis. Lancet Gastroenterol Hepatol 2023;8:704-713.

**Table 1. Clinical characteristics of the participants**

| | Reference Range | MASH (n = 78) | MAS (n = 156) | VH (n = 78) | NL (n = 156) | *P* (MASH v.s. Controls) | | | *P** | *P*** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | MAS | VH | NL | MAS-NL | VH-NL |
| Male, n (%) | / | 40 (52.63) | 80 (52.63) | 40 (52.63) | 80 (52.63) | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Age, years | / | 59.37 ± 7.12 | 59.36 ± 7.13 | 58.09 ± 7.41 | 59.37 ± 7.10 | 0.392 | 0.998 | 0.434 | 0.987 | 0.789 |
| BMI, kg/m$^2$ | 18.5 - 30 | 33.51 ± 5.78 | 31.23 ± 5.71 | 27.72 ± 5.49 | 27.43 ± 4.71 | 0.423 | **<0.001** | **<0.001** | **<0.001** | **0.004** |
| ALT, U/L | 0 - 40 | 37.26 (53.55, 25.43) | 29.23 (41.14, 20.15) | 20.46 (16.42, 30.47) | 19.75 (16.19, 25.77) | 0.767 | **<0.001** | **<0.001** | **<0.001** | 0.014 |
| AST, U/L | 0 - 40 | 37.00 (27.60, 50.40) | 27.80 (23.20, 36.30) | 25.10 (21.20, 29.83) | 25.20 (22.20, 28.20) | 0.733 | **0.001** | **<0.001** | **<0.001** | 0.036 |
| FIB-4 | < 3.25 | 2.12 ± 1.31 | 1.46 ± 0.89 | 1.63 ± 0.94 | 1.47 ± 0.84 | **<0.001** | **0.010** | **<0.001** | 0.938 | 0.214 |
| ALB, g/L | 40 – 55 | 44.21 ± 3.26 | 44.83 ± 2.65 | 44.02 ± 2.68 | 45.3 ± 2.6 | 0.152 | 0.903 | 0.218 | 0.144 | 0.013 |
| ALP, U/L | 40 – 160 | 101.81 ± 34.06 | 91.1 ± 40.28 | 90.28 ± 39.96 | 83.42 ± 20.92 | 0.142 | 0.186 | **0.005** | 0.040 | 0.012 |
| GGT, U/L | 10 - 60 | 78.90 (42.50, 179.75) | 46.70 (27.60, 70.00) | 26.30 (19.70, 43.50) | 27.10 (21.40, 41.70) | **0.002** | 0.058 | **<0.001** | **<0.001** | 0.486 |
| PLT, ×10$^9$/L | 100 - 350 | 217.71 ± 75.23 | 256.65 ± 69.57 | 241.34 ± 58.84 | 250.96 ± 60.99 | 0.029 | 0.511 | 0.018 | 0.450 | 0.282 |
| TBiL, μmol/L | 1.7 - 17.1 | 9.74 ± 5.32 | 9.28 ± 4.35 | 9.32 ± 5.8 | 9.18 ± 6.03 | 0.769 | 0.680 | 0.326 | 0.869 | 0.374 |
| FPG, mmol/L | 3.9 - 6.1 | 7.06 ± 4.3 | 5.97 ± 2.99 | 5.37 ± 1.91 | 5.15 ± 1.43 | 0.404 | **0.004** | 0.016 | **0.004** | 0.519 |
| CRP, mg/L | < 8 | 3.51 (1.60, 5.73) | 2.35 (1.09, 5.25) | 1.81 (0.52, 3.71) | 1.37 (0.72, 2.93) | 0.276 | **0.004** | **<0.001** | 0.016 | **0.007** |
| TC, mmol/L | 2.84 - 5.68 | 4.81 ± 1.31 | 5.44 ± 1.26 | 5.31 ± 1.02 | 5.8 ± 1.24 | 0.910 | **<0.001** | **<0.001** | 0.012 | 0.108 |
| TG, mmol/L | 0.56 - 1.70 | 2.19 ± 1.13 | 2.32 ± 1.42 | 1.6 ± 0.85 | 1.83 ± 1.01 | 0.532 | 0.015 | **<0.001** | **<0.001** | 0.862 |
| HDL, mmol/L | 1.14 - 1.91 | 1.16 ± 0.33 | 1.29 ± 0.37 | 1.41 ± 0.35 | 1.43 ± 0.37 | 0.295 | 0.016 | **0.003** | **0.003** | 0.056 |
| LDL, mmol/L | 1 - 3 | 2.99 ± 1.00 | 3.37 ± 0.92 | 3.31 ± 0.78 | 3.66 ± 0.95 | 0.865 | 0.439 | 0.178 | **0.009** | 0.183 |
| UA, μmol/L | 90 - 360 | 351.37 ± 107.82 | 348.15 ± 91.19 | 330.4 ± 105.81 | 324.26 ± 85.09 | 0.272 | 0.981 | 0.430 | 0.021 | 0.543 |
| Urea, mmol/L | 1.8 – 7.1 | 6.16 ± 3.21 | 5.63 ± 1.54 | 6.05 ± 2.63 | 5.73 ± 1.44 | 0.211 | 0.127 | 0.014 | 0.581 | 0.903 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Comorbidities at baseline | | | | | | | | | |
| Diabetes | 5 (6.41) | 4 (2.56) | 8 (10.26) | 20 (12.82) | 0.279 | 0.385 | 0.135 | **<0.001** | 0.569 |
| Hypertension | 9 (11.54) | 26 (16.67) | 19 (24.36) | 34 (21.79) | 0.400 | 0.060 | 0.084 | 0.315 | 0.783 |
| Digestive cancers | 3 (3.85) | 3 (1.92) | 2 (2.56) | 6 (3.85) | 0.661 | 1.000 | 1.000 | 1.000 | 0.899 |
| GIT/BT diseases | 10 (12.82) | 25 (16.03) | 19 (24.36) | 44 (28.21) | 0.650 | 0.100 | 0.014 | 0.014 | 0.639 |

Note: Data are presented as means± SDs or frequencies (percentages). $P < 0.01$ (Bonferroni adjusted α) is considered statistically significant. GIT/BT, gastrointestinal tract and biliary tract.

**Table 2. Levels of proteins in MASH and three control groups**

| | **Detected range in blood by PEA** | **MASH** (n = 78) | **MAS** (n = 156) | **VH** (n = 78) | **NL** (n = 156) |
|---|---|---|---|---|---|
| **CDCP1,** NPX | 1.0-4.0 | 0.91 ± 0.88 | 0.39 ± 0.74 | 0.36 ± 0.90 | 0.11 ± 0.62 |
| **FGF21,** NPX | 3.0-8.0 | 1.63 ± 1.65 | 1.18 ± 1.47 | -0.18 ± 1.42 | 0.02 ± 1.33 |
| **IL-6,** NPX | 2.0-6.0 | 1.19 ± 1.11 | 0.55 ± 0.79 | 0.30 ± 0.88 | 0.19 ± 0.97 |
| **FABP4,** NPX | 3.0-6.0 | 0.86 ± 0.92 | 0.54 ± 0.86 | 0.11 ± 0.82 | 0.03 ± 0.75 |
| **GDF15,** NPX | 3.0-5.0 | 1.15 ± 0.91 | 0.45 ± 0.71 | 0.37 ± 0.70 | 0.09 ± 0.51 |
| **THBS2,** NPX | 4.0-6.0 | 1.09 ± 1.06 | 0.37 ± 0.67 | 0.21 ± 0.61 | -0.01 ± 0.31 |

Note: NPX, normalized protein expression; PEA, proximity extension assay; The detected range of six circulating protein biomarkers in blood is obtained from The Human Protein Atlas.

## Figure Legends

**Figure 1. Flow chart of participant selection and mean time to diagnosis of liver diseases.** Diagnosis time was defined as the lag time from recruitment (baseline) to disease diagnosis. Each black dot represents a single patient.

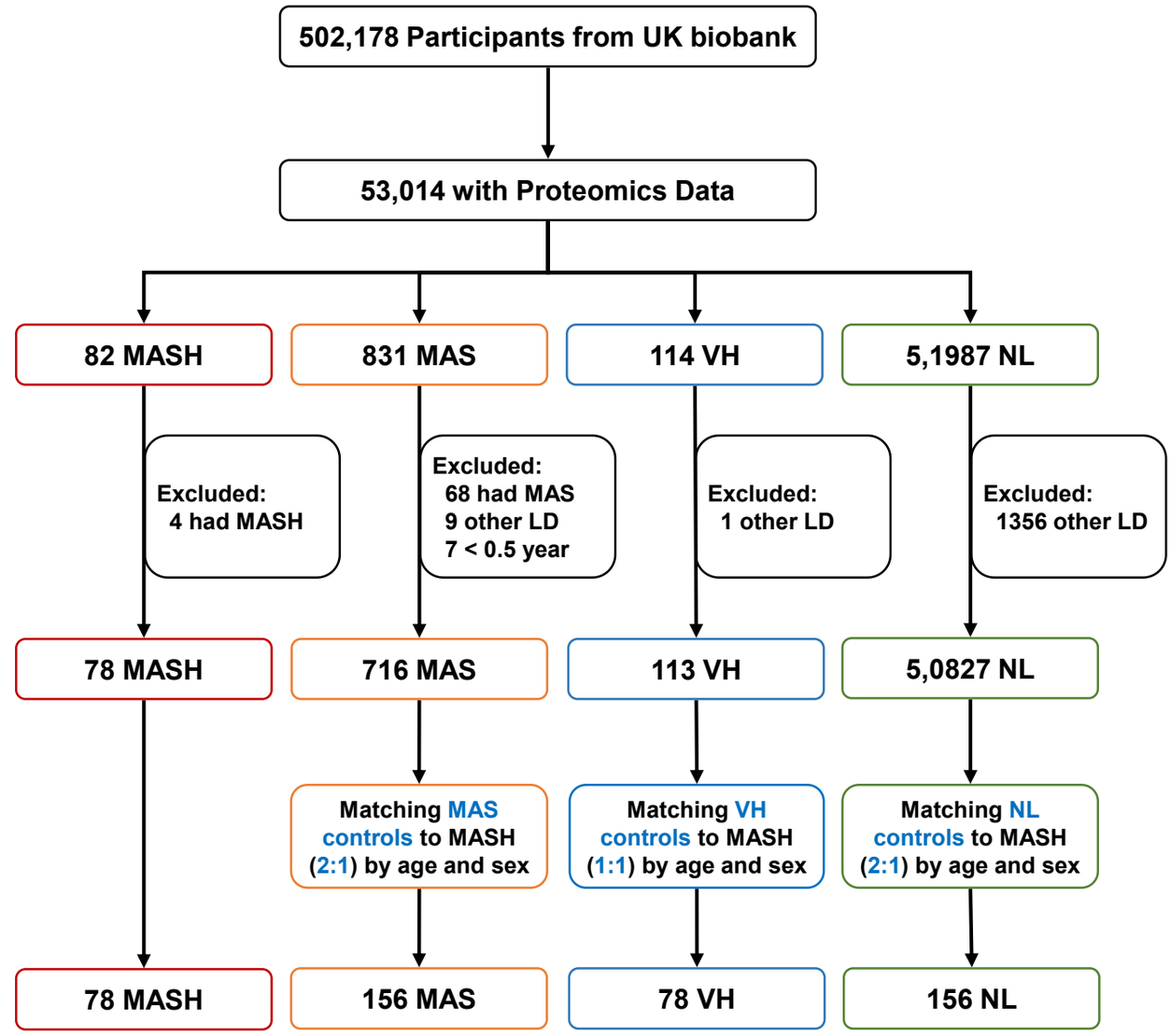
**Figure 2. Violin plots comparing baseline levels of the six proteomic biomarkers among MASH, MAS, VH and NL control groups.** The expression levels of the six biomarkers were expressed as normalized protein expression (NPX) units. P value ranges were indicated with *s for comparisons among different groups. ns, non-significant; *; P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001. *P* < 0.01 (Bonferroni adjusted α) is considered statistically significant.

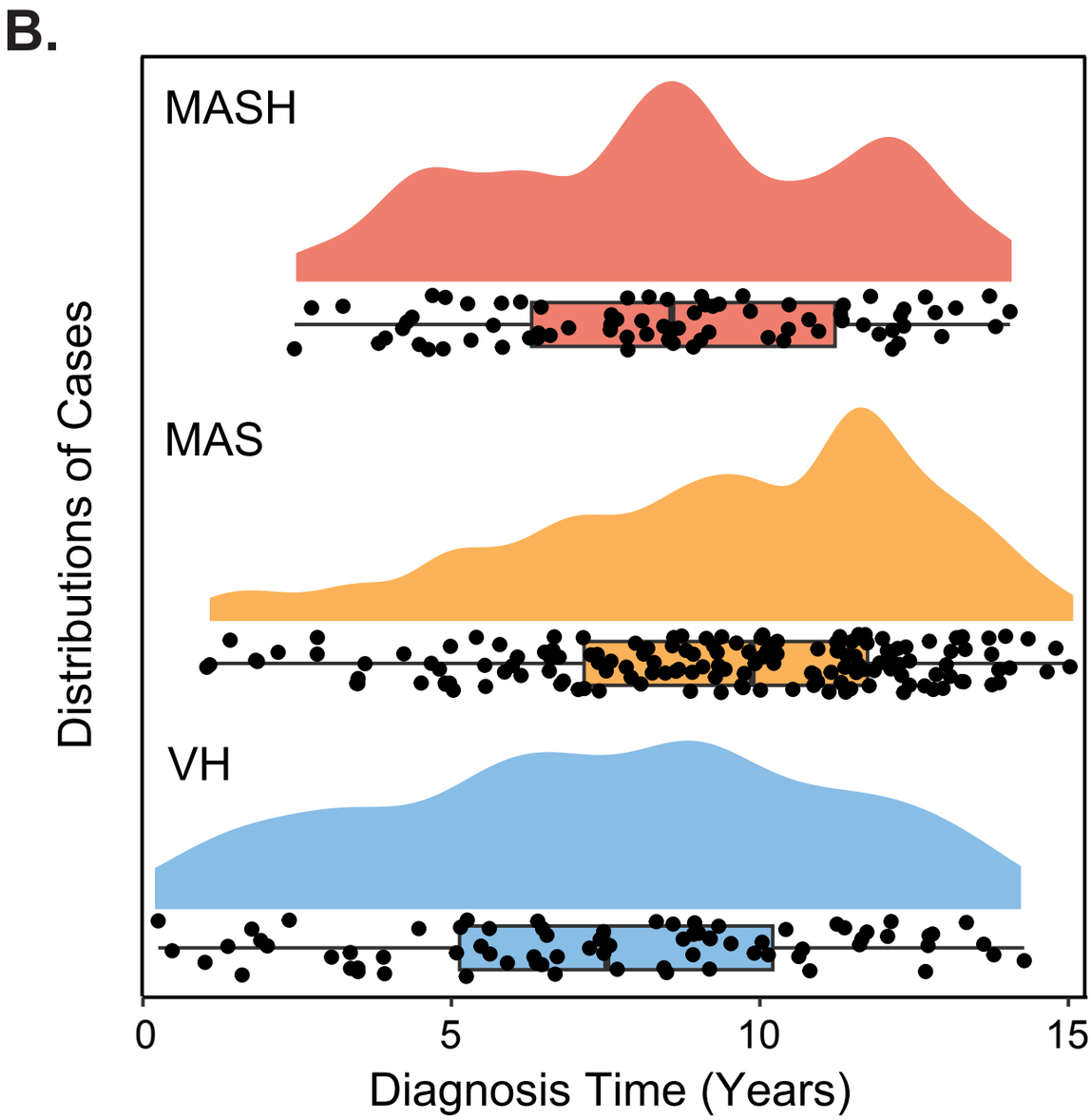
**Figure 3. Spearman correlations among baseline levels of the six circulating biomarkers and clinical characteristics**. Statistically significant correlations are shown with numbers, while the non-significant correlation coefficients are blank in the boxes.

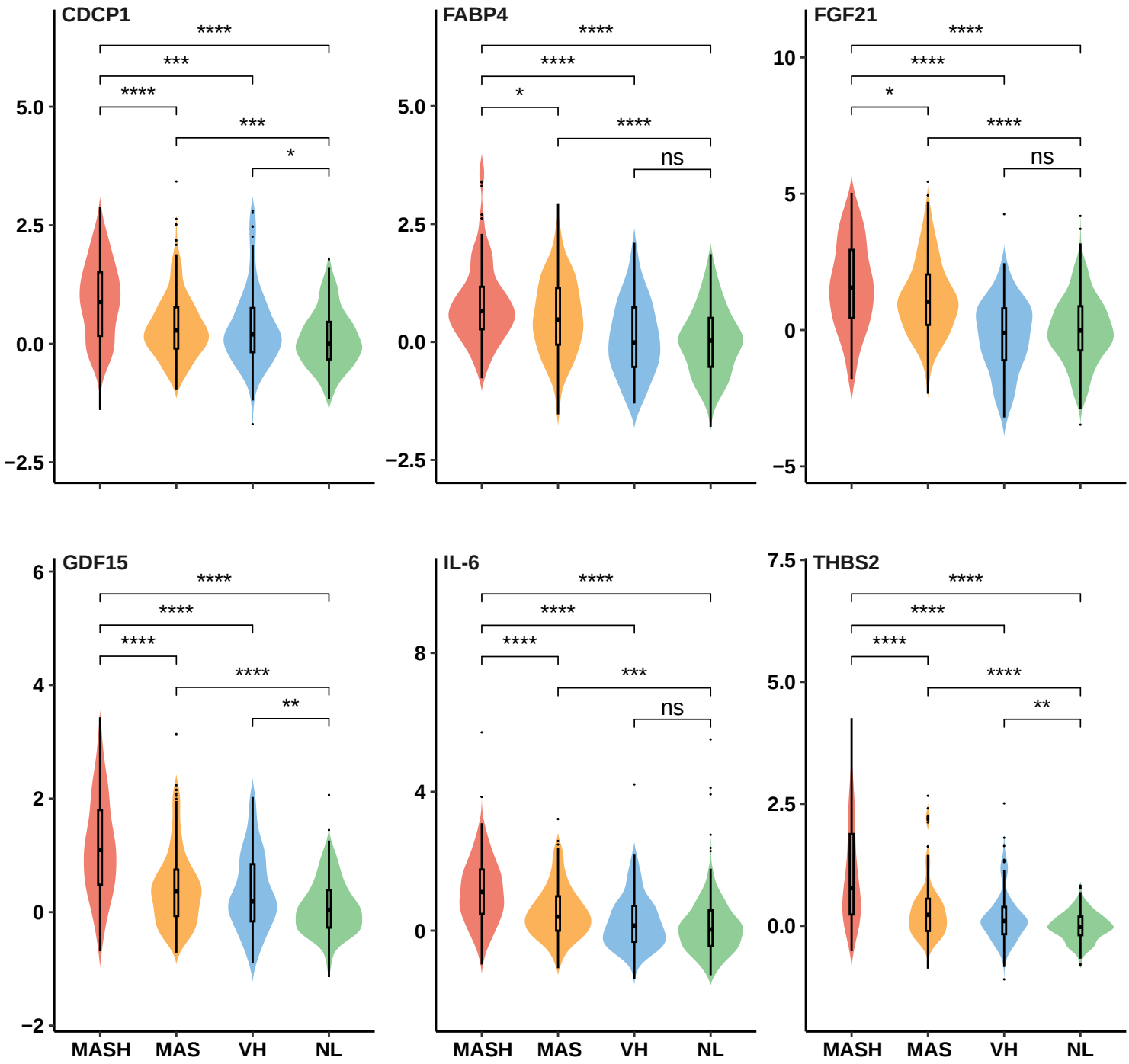
**Figure 4. Forest plot of risk effects of the six circulating protein biomarkers for MASH, MAS and VH development.** Shown are forest plots of HRs adjusted by three models in predicting MASH against NL (A), MAS against NL (B), and VH against NL (C).

**Figure 5. Time-dependent ROC curves of the six biomarkers, two combinatory models, and clinical factors for MASH prediction against NL controls.**

**Figure 6. Time-dependent ROC curves of the six biomarkers, two combinatory models, and clinical factors for MASH prediction against MAS.**


A.
502,178 Participants from UK biobank
53,014 with Proteomics Data
82 MASH
831 MAS
114 VH
5,1987 NL
Excluded:
4 had MASH
Excluded:
68 had MAS
9 other LD
7 < 0.5 year
Excluded:
1 other LD
Excluded:
1356 other LD
78 MASH
716 MAS
113 VH
5,0827 NL
Matching MAS controls to MASH (2:1) by age and sex
Matching VH controls to MASH (1:1) by age and sex
Matching NL controls to MASH (2:1) by age and sex
78 MASH
156 MAS
78 VH
156 NL


B.
MASH
MAS
VH
Distributions of Cases
0
5
10
15
Diagnosis Time (Years)

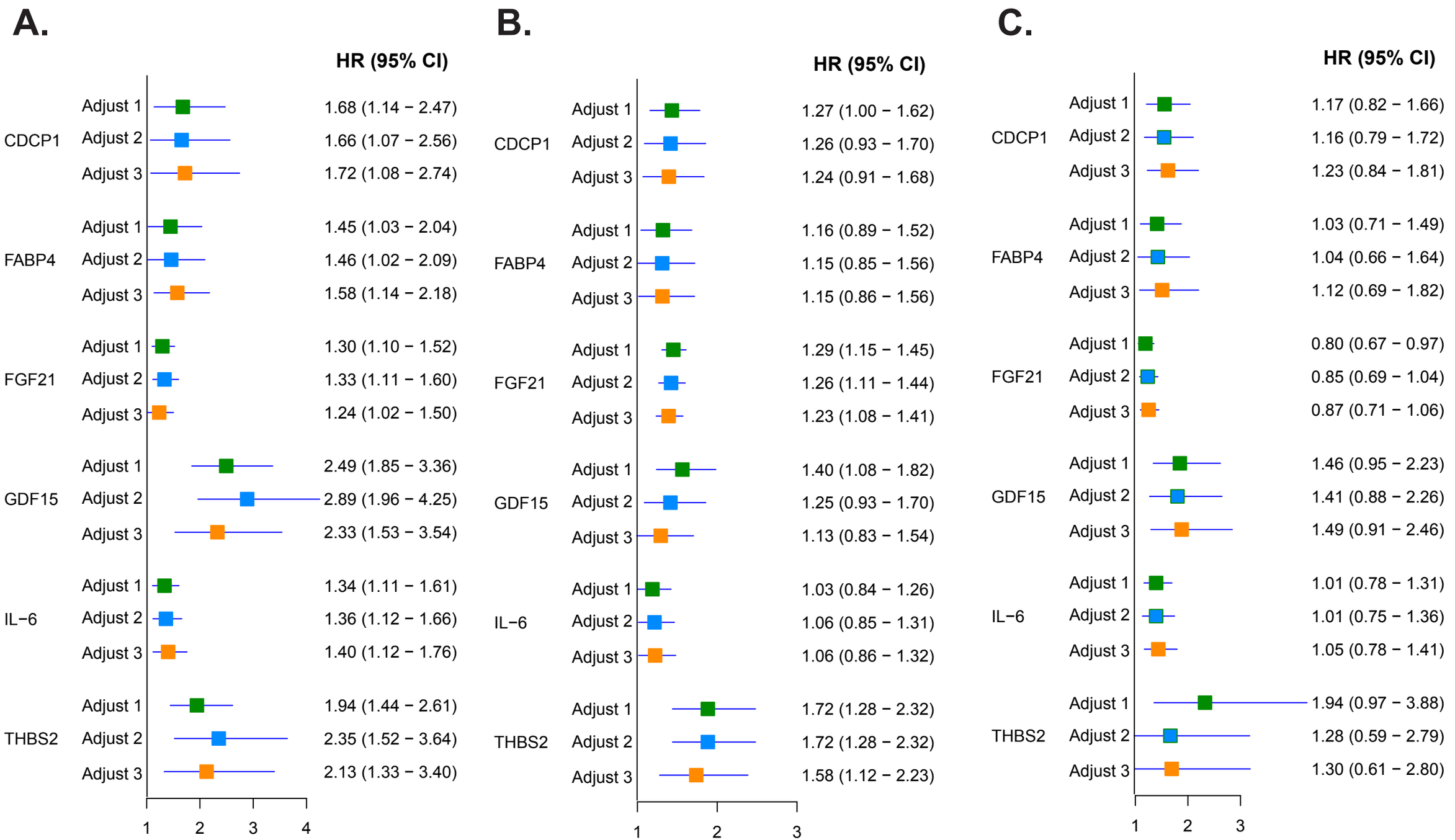

CDCP1
FABP4
FGF21
GDF15
IL-6
THBS2
MASH
MAS
VH
NL
****
***
**
*
ns

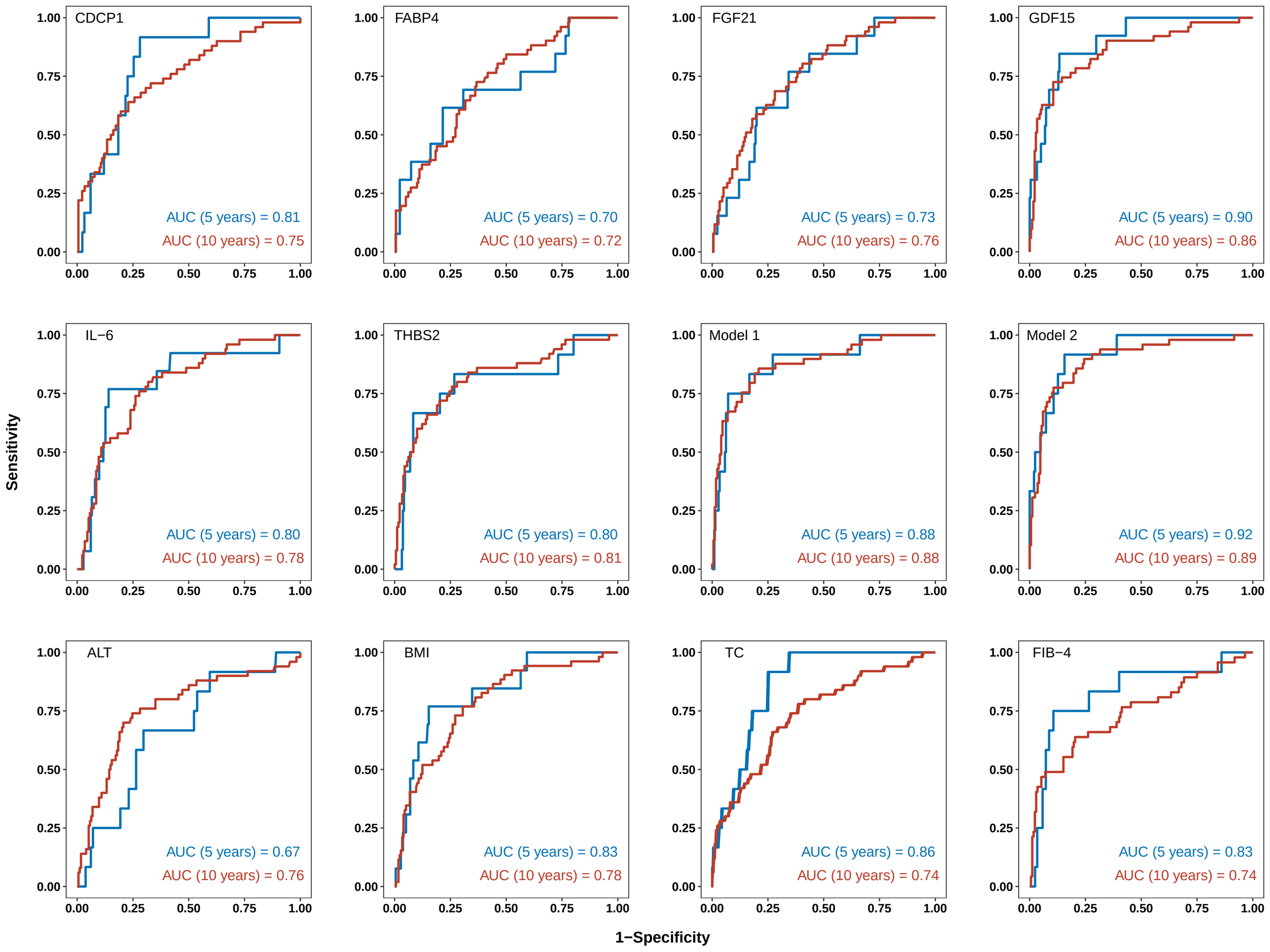
A.

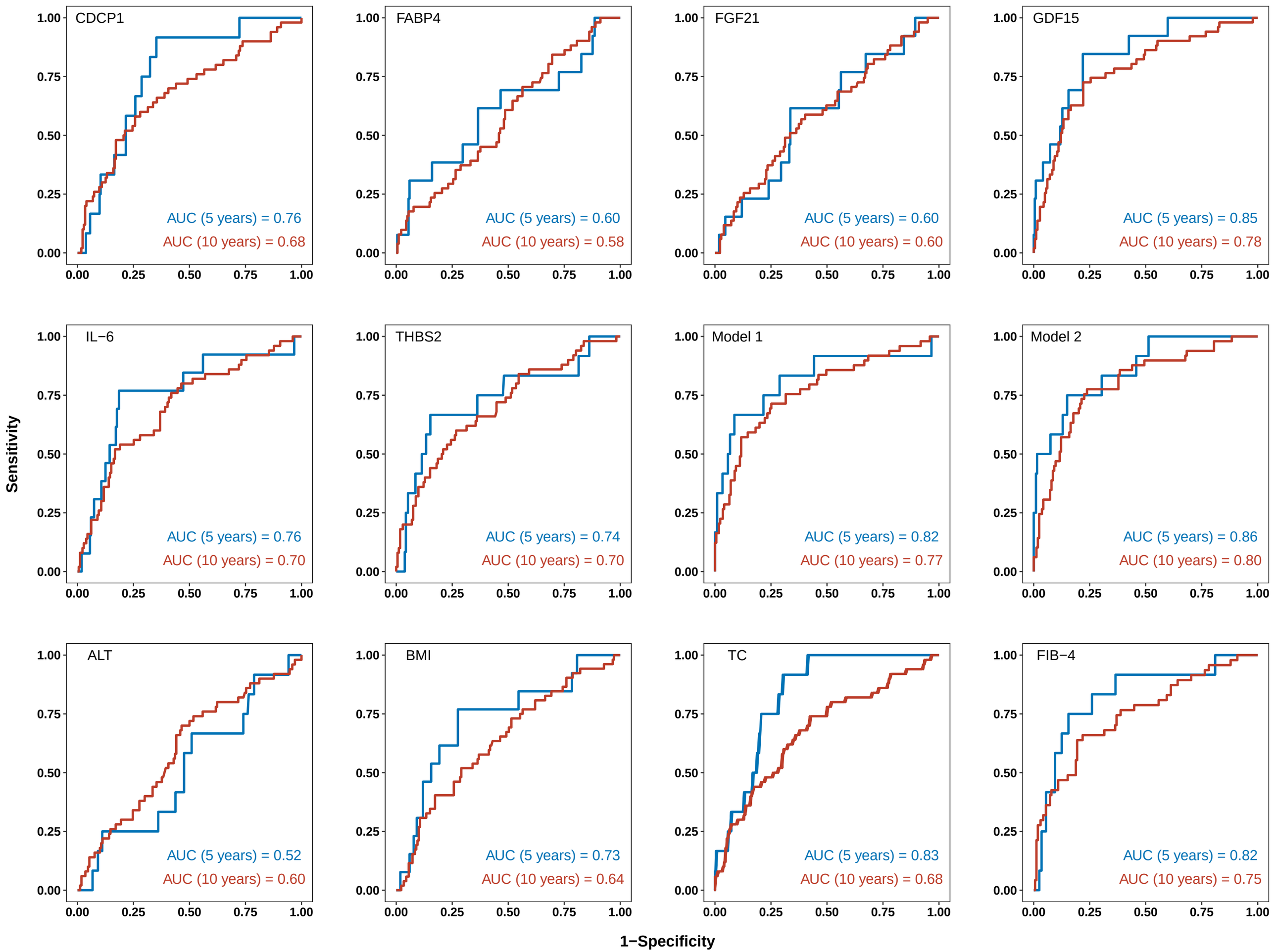
| | CDCP1 | FABP4 | FGF21 | GDF15 | IL-6 | THBS2 | ALT | AST | ALP | ALB | GGT |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CDCP1 | | 0.23 | 0.33 | 0.59 | 0.40 | 0.47 | 0.39 | 0.4 | 0.3 | −0.17 | 0.4 |
| FABP4 | | | 0.37 | 0.37 | 0.38 | 0.28 | 0.18 | | 0.22 | | 0.21 |
| FGF21 | | | | 0.41 | 0.41 | 0.40 | 0.48 | 0.41 | 0.16 | | 0.48 |
| GDF15 | | | | | 0.49 | 0.50 | 0.3 | 0.3 | 0.21 | −0.19 | 0.34 |
| IL-6 | | | | | | 0.40 | 0.25 | 0.24 | 0.3 | −0.23 | 0.33 |
| THBS2 | | | | | | | 0.45 | 0.41 | 0.21 | −0.18 | 0.5 |

B.

| | FIB-4 | BMI | FPG | CRP | Uric Acid | Urate | PLT | TC | TG | HDL | LDL |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CDCP1 | 0.38 | 0.24 | 0.28 | 0.33 | | | −0.27 | −0.11 | 0.25 | −0.27 | |
| FABP4 | | 0.63 | 0.17 | 0.36 | 0.29 | 0.12 | 0.13 | | 0.32 | −0.17 | |
| FGF21 | | 0.43 | 0.31 | 0.27 | 0.36 | −0.12 | | −0.15 | 0.44 | −0.36 | −0.14 |
| GDF15 | 0.41 | 0.34 | 0.3 | 0.26 | 0.21 | 0.22 | −0.28 | −0.29 | 0.28 | −0.36 | −0.27 |
| IL-6 | 0.2 | 0.43 | 0.18 | 0.52 | 0.25 | | −0.14 | −0.16 | 0.21 | −0.28 | −0.13 |
| THBS2 | 0.24 | 0.36 | 0.17 | 0.33 | | | −0.2 | −0.14 | 0.24 | −0.27 | −0.13 |


A.
HR (95% CI)
CDCP1
Adjust 1
1.68 (1.14 – 2.47)
Adjust 2
1.66 (1.07 – 2.56)
Adjust 3
1.72 (1.08 – 2.74)
FABP4
Adjust 1
1.45 (1.03 – 2.04)
Adjust 2
1.46 (1.02 – 2.09)
Adjust 3
1.58 (1.14 – 2.18)
FGF21
Adjust 1
1.30 (1.10 – 1.52)
Adjust 2
1.33 (1.11 – 1.60)
Adjust 3
1.24 (1.02 – 1.50)
GDF15
Adjust 1
2.49 (1.85 – 3.36)
Adjust 2
2.89 (1.96 – 4.25)
Adjust 3
2.33 (1.53 – 3.54)
IL−6
Adjust 1
1.34 (1.11 – 1.61)
Adjust 2
1.36 (1.12 – 1.66)
Adjust 3
1.40 (1.12 – 1.76)
THBS2
Adjust 1
1.94 (1.44 – 2.61)
Adjust 2
2.35 (1.52 – 3.64)
Adjust 3
2.13 (1.33 – 3.40)
1
2
3
4
B.
HR (95% CI)
CDCP1
Adjust 1
1.27 (1.00 – 1.62)
Adjust 2
1.26 (0.93 – 1.70)
Adjust 3
1.24 (0.91 – 1.68)
FABP4
Adjust 1
1.16 (0.89 – 1.52)
Adjust 2
1.15 (0.85 – 1.56)
Adjust 3
1.15 (0.86 – 1.56)
FGF21
Adjust 1
1.29 (1.15 – 1.45)
Adjust 2
1.26 (1.11 – 1.44)
Adjust 3
1.23 (1.08 – 1.41)
GDF15
Adjust 1
1.40 (1.08 – 1.82)
Adjust 2
1.25 (0.93 – 1.70)
Adjust 3
1.13 (0.83 – 1.54)
IL−6
Adjust 1
1.03 (0.84 – 1.26)
Adjust 2
1.06 (0.85 – 1.31)
Adjust 3
1.06 (0.86 – 1.32)
THBS2
Adjust 1
1.72 (1.28 – 2.32)
Adjust 2
1.72 (1.28 – 2.32)
Adjust 3
1.58 (1.12 – 2.23)
1
2
3
C.
HR (95% CI)
CDCP1
Adjust 1
1.17 (0.82 – 1.66)
Adjust 2
1.16 (0.79 – 1.72)
Adjust 3
1.23 (0.84 – 1.81)
FABP4
Adjust 1
1.03 (0.71 – 1.49)
Adjust 2
1.04 (0.66 – 1.64)
Adjust 3
1.12 (0.69 – 1.82)
FGF21
Adjust 1
0.80 (0.67 – 0.97)
Adjust 2
0.85 (0.69 – 1.04)
Adjust 3
0.87 (0.71 – 1.06)
GDF15
Adjust 1
1.46 (0.95 – 2.23)
Adjust 2
1.41 (0.88 – 2.26)
Adjust 3
1.49 (0.91 – 2.46)
IL−6
Adjust 1
1.01 (0.78 – 1.31)
Adjust 2
1.01 (0.75 – 1.36)
Adjust 3
1.05 (0.78 – 1.41)
THBS2
Adjust 1
1.94 (0.97 – 3.88)
Adjust 2
1.28 (0.59 – 2.79)
Adjust 3
1.30 (0.61 – 2.80)
1
2
3


CDCP1
AUC (5 years) = 0.81
AUC (10 years) = 0.75
FABP4
AUC (5 years) = 0.70
AUC (10 years) = 0.72
FGF21
AUC (5 years) = 0.73
AUC (10 years) = 0.76
GDF15
AUC (5 years) = 0.90
AUC (10 years) = 0.86
IL−6
AUC (5 years) = 0.80
AUC (10 years) = 0.78
THBS2
AUC (5 years) = 0.80
AUC (10 years) = 0.81
Model 1
AUC (5 years) = 0.88
AUC (10 years) = 0.88
Model 2
AUC (5 years) = 0.92
AUC (10 years) = 0.89
ALT
AUC (5 years) = 0.67
AUC (10 years) = 0.76
BMI
AUC (5 years) = 0.83
AUC (10 years) = 0.78
TC
AUC (5 years) = 0.86
AUC (10 years) = 0.74
FIB−4
AUC (5 years) = 0.83
AUC (10 years) = 0.74
Sensitivity
1−Specificity
0.00
0.25
0.50
0.75
1.00


CDCP1
AUC (5 years) = 0.76
AUC (10 years) = 0.68
FABP4
AUC (5 years) = 0.60
AUC (10 years) = 0.58
FGF21
AUC (5 years) = 0.60
AUC (10 years) = 0.60
GDF15
AUC (5 years) = 0.85
AUC (10 years) = 0.78
IL−6
AUC (5 years) = 0.76
AUC (10 years) = 0.70
THBS2
AUC (5 years) = 0.74
AUC (10 years) = 0.70
Model 1
AUC (5 years) = 0.82
AUC (10 years) = 0.77
Model 2
AUC (5 years) = 0.86
AUC (10 years) = 0.80
ALT
AUC (5 years) = 0.52
AUC (10 years) = 0.60
BMI
AUC (5 years) = 0.73
AUC (10 years) = 0.64
TC
AUC (5 years) = 0.83
AUC (10 years) = 0.68
FIB−4
AUC (5 years) = 0.82
AUC (10 years) = 0.75
Sensitivity
1−Specificity
0.00
0.25
0.50
0.75
1.00