

1 Iron biomarkers predict peripheral artery disease in females: a cross-sectional 2 analysis of HEIST-DiC and NHANES participants

3
4 Anand Ruban Agarvas^{1,2}, Stefan Kopf^{1,3,4}, Tiago J.S. Lopes^{1,2,10}, Paul Thalmann^{1,5}, José
5 Manuel Fernández-Real⁶, Peter Nawroth^{1,3}, *Martina U. Muckenthaler^{1,2,7,8,9}

6

7 Affiliations

8 ¹Heidelberg University, Heidelberg, Germany.

9 ²Center For Translational Biomedical Iron Research, Department of Pediatric Hematology,
10 Oncology Immunology and Pulmonology, University Hospital Heidelberg, Heidelberg,
11 Germany.

12 ³Clinic for Endocrinology, Diabetology, Metabolic Diseases and Clinical Chemistry (Internal
13 Medicine 1), University Hospital Heidelberg, Heidelberg, Germany.

14 ⁴German Center of Diabetes Research (DZD), Neuherberg, Germany.

15 ⁵Institute of Medical Biometry, University Hospital Heidelberg, Heidelberg, Germany.

16 ⁶Department of Diabetes, Endocrinology and Nutrition, Institut d'Investigació Biomèdica de
17 Girona (IdIBGi), CIBEROBN (CB06/03/010) and Instituto de Salud Carlos III (ISCIII), Girona
18 17007, Spain.

19 ⁷Molecular Medicine Partnership Unit, Heidelberg, Germany.

20 ⁸Translational Lung Research Center Heidelberg (TLRC), German Center for Lung
21 Research (DZL), Heidelberg, Germany.

22 ⁹German Centre for Cardiovascular Research (DZHK), Partner Site Heidelberg/Mannheim,
23 Germany.

24 ¹⁰Nezu Life Sciences, Karlsruhe, Germany

25

26 *Corresponding author

27

28

29 Address for correspondence

30 Department of Pediatric Hematology, Oncology and Immunology, Heidelberg University,
31 Heidelberg, Germany.
32 Tel.: +496221566923; Fax: +496221564580. E-mail: [martina.muckenthaler@med.uni-](mailto:martina.muckenthaler@med.uni-heidelberg.de)
33 [heidelberg.de](mailto:martina.muckenthaler@med.uni-heidelberg.de) (M.U. Muckenthaler).

34 Abstract

35 Background

36 Iron overload has been proposed as a risk factor for atherosclerosis but the available data
37 are controversial. Here we investigated whether iron biomarkers within physiological limits
38 that frequently differ between sexes are associated with peripheral arterial disease (PAD)
39

40 Methods

41 Using two different analytical approaches (machine-learning and logistic regression), we
42 studied the association between blood iron biomarkers and PAD in 368 individuals from the
43 Heidelberg Study on Diabetes and Complications (HEIST-DiC) and in 5101 individuals from
44 the National Health and Nutrition Examination Survey (NHANES 1999-2004).
45

46 Results

47 We found that iron biomarkers were among the top predictors of PAD in the machine-
48 learning classification in both cohorts. In the HEIST-DiC cohort, ferritin, iron, and transferrin
49 were ranked among the top predictive markers, while in the NHANES cohort ferritin, total
50 iron binding capacity (TIBC), transferrin saturation (TSAT) and iron showed high predictive
51 power. In the regression analysis ferritin showed a positive interaction among females in the
52 HEIST-DiC (OR 2.68, 95% CI 0.94-7.61, $p=0.057$) and NHANES cohorts (OR 1.76, 95% CI
53 1.16-2.67, $p=0.008$). The multivariable regression analysis of the NHANES cohort, detected
54 a nonlinear relationship between ferritin and PAD in that certain ferritin ranges (48-97 ng/mL:
55 OR 14.59, 95% CI 1.6-135.93, $p=0.019$; 98-169 ng/mL: OR 171.07, 95% CI 1.27-23404, $p=$
56 0.039) in females associated positively with PAD.
57

58 Conclusion

59 Taken together our data show that iron biomarkers, importantly, elevated ferritin is
60 associated with clinically apparent PAD in females. These findings add to the body of
61 evidence suggesting sex differences in PAD and highlight a possible role of iron (directly or
62 indirectly) in this relationship.
63

64 Keywords

65 Iron, Vascular Diseases, Arterial Occlusive Diseases, Atherosclerosis, Peripheral Arterial
66 Disease
67

68 Introduction

69 Peripheral arterial disease (PAD) is a common atherosclerotic vascular disorder of the
70 arteries of the lower limbs that affects millions of individuals worldwide^{1,2}. Compared to
71 cardiovascular and cerebrovascular disease, PAD is less well studied, and our
72 understanding of the disease process is still poor. The known risk factors of PAD include old
73 age, smoking, and comorbidities such as diabetes, hypertension or Vitamin D excess¹.

74

75 Iron is an essential micronutrient, however, its accumulation in excess contributes to the
76 generation of reactive oxygen species. Various regulatory systems maintain iron
77 homeostasis by controlling absorption, transport, and storage³. Under physiological
78 conditions, iron is bound to transferrin in the plasma, and transported to cells that require
79 iron. Usually, less than one-third of transferrin's binding sites are occupied by iron, allowing
80 free iron entering the circulation to be efficiently sequestered. However, in pathological
81 states associated with iron overload, the binding capacity of transferrin is nearly saturated
82 resulting in the appearance of 'non-transferrin-bound iron (NTBI)'. Therefore, iron overload is
83 discussed as a risk factor for the development of atherosclerosis⁴. Accordingly, inhibition of
84 ferroptosis, an iron-dependent form of cell death, has been shown to protect from
85 atherosclerosis⁵. Iron and PAD may be linked by sex hormones. Changes in estrogen and
86 progesterone levels associated with menopause were shown to impair endothelial function,
87 especially among women^{6,7}. Both hormones further control hepcidin, the master regulator of
88 iron homeostasis in opposing directions, with estrogen⁸ inhibiting and progesterone⁹
89 activating hepcidin expression.

90

91 In our previous work, we demonstrated that iron overload in mouse models and humans with
92 hemochromatosis show an aggravation of atherosclerosis and vessel damage,
93 respectively¹⁰. Specifically, we found that NTBI in the circulation contributes to
94 atherosclerosis by oxidising low-density lipoproteins (LDL) and by inducing vascular
95 dysfunction; in addition, iron deposition in the vasculature of hemochromatotic mice was
96 associated with more severe atherosclerosis¹⁰. Whether or not iron levels within
97 physiological limits affect atherogenesis is currently unclear and available data are
98 controversial. Most studies undertaken so far have investigated the relationships between
99 iron-related parameters [e.g. plasma iron, ferritin, transferrin, transferrin saturation (TSAT)]
100 and cardiovascular disease¹¹⁻¹⁷ but associations with PAD are rarely addressed.

101

102 To address this, we analyzed iron biomarkers for their associations with PAD and by
103 considering sex-specific differences. We used two different datasets and orthogonal
104 approaches: 1) machine learning (ML) classification and 2) logistic regression. Both methods

can be used for analyzing relationships between variables, but they differ in their approach and application. ML classification tasks can analyze complex patterns in multidimensional data and identify nonlinear interactions among variables. They are particularly useful when manual specification of relationships is challenging as is the case for human clinical data. In contrast, regression analysis quantifies the strength and direction of associations between variables and provides interpretable results in terms of odds ratios. Both approaches carry their strengths and limitations. However, we expect that true associations in the data (in this case, sex-specific differential associations between iron biomarkers and PAD) should be detected by both methods. Therefore, we initially used the ML approach to provide unbiased insight into our data and subsequently regression analyses to validate identified relationship. An outline of the study is shown in Figure 1.

116

117 **Methods**

118 ***HEIST-DiC study***

The Heidelberg Study on Diabetes and Complications (HEIST-DiC; <https://clinicaltrials.gov/NCT03022721>) is a prospective, observational study designed to analyze the development of complications in individuals with diabetes and prediabetes. We analyzed the data from 368 individuals enrolled in the HEIST-DiC study. The selection criteria for participants have been described previously¹⁸. We excluded participants with values of serum iron biomarkers suggestive of hemochromatosis¹⁹. In addition, since ferritin not only indicates iron stores, but also is an acute phase protein that responds to inflammation, we excluded participants with ferritin > 400 ng/mL. For all participants, a detailed medical history was obtained. The ankle-brachial index (ABI) was measured using noninvasive blood pressure measurements of arms and ankles (ABI System 1000; Boso d.o.o.). Blood samples were obtained in the fasting state and all baseline parameters were analyzed under standardized conditions in the central laboratory of the University Hospital of Heidelberg. Serum samples were stored at -80°C until further analysis. We analyzed iron, ferritin, and TSAT as biomarkers of systemic iron status, and assayed serum intercellular adhesion molecule-1 (ICAM1) and vascular-cell adhesion molecule-1 (VCAM1) as markers of vascular dysfunction. We measured ICAM1 (#171B6009M) and VCAM1 (#171B6022M) in serum using multiplex magnetic bead-array-based technology on a BioPlex200 system (Bio-Rad) according to the manufacturer's instructions. All participants provided written informed consent, and the ethics committees of the University of Heidelberg approved this study (Decision No. 204/2004, 400/2010, and S-383/2016) according to the Declaration of Helsinki.

139 ***NHANES***

140 The National Health and Nutrition Examination Survey (NHANES) is designed to assess the
141 health and nutritional status of adults and children in the United States. The survey includes
142 both an interview (in which participants are asked questions about their health and lifestyle),
143 and a physical examination component (in which various measurements and samples are
144 collected). We obtained data from 15,969 participants over 40 years in 3 consecutive cycles
145 of NHANES (1999-2004) for whom ABI was measured.

146

147 The demographic variables, age (in years) and body mass index (BMI) were used as is.
148 Smoking and alcohol were categorized according to a previous study²⁰. We further
149 categorized diabetes based on self-reported history, current use of insulin or oral
150 hypoglycemic agents, fasting plasma glucose (FPG) ≥ 7.0 mM, or glycated hemoglobin
151 (HbA1c) $\geq 6.5\%$. Prediabetes was defined among those with an absent self-reported history
152 of diabetes by either FPG 5.6-6.9 mM or HbA1c 5.7-6.4%. We categorized hypertension
153 based on self-reported history or current use of anti-hypertensive agents or systolic blood
154 pressure >130 mmHg or diastolic blood pressure >80 mmHg. The presence of nonalcoholic
155 fatty liver disease (NAFLD) was categorized by calculating the Fibrosis-4 Index score (FIB-4
156 score >3.25)²¹. As an index of insulin resistance, we calculated HOMA2-IR using C-peptide
157 values instead of insulin²². The use of antiplatelet and antihyperlipidemic medications was
158 categorized based on medication intake.

159

160 We applied the following exclusion criteria: not even one ABI measurement available, likely
161 hemochromatosis¹⁹, anemia (hemoglobin <10 g/L), elevated ferritin (>400 ng/mL), cancer, on
162 current hormonal medications or iron supplements, donated blood in the last 1 month, with
163 features of kidney disease (on dialysis or severe albuminuria), pregnant or breastfeeding in
164 the last year, high risk of liver fibrosis, concurrent inflammation [C-Reactive Protein (CRP)
165 >3 mg/dL] and with a feature of familial hyperlipidemia [LDL-Cholesterol (LDLc) >4.9 mM].
166 This resulted in a final sample number of 5101 participants. We obtained data from the
167 following iron biomarkers: iron, ferritin, total iron binding capacity (TIBC), and TSAT. Assays
168 for the analytes in NHANES were performed under strict laboratory protocols (for more
169 details on data fields and procedure manuals, see Supplementary Table 1).

170

171 *Categorization as PAD*

172 In both cohorts, ABI values were measured on both the right and left limbs. We categorized
173 PAD when the ABI value on one of the sides was <0.9 or >1.4 as per recommendations²³.

174

175 **Implementation of the ML framework**

176

177 First, we preprocessed the HEIST-DiC and the NHANES dataset and removed all features
178 that contained more than 10% of missing values. Additionally, to made both cohorts
179 comparable by only considering the cases that did not have missing values for ferritin,
180 diabetes, CRP and high blood pressure in the NHANES cohort. For the HEIST-DiC dataset,
181 we have used the leave-one-out split strategy, and for each fold, we filled in the missing
182 values using the average of each column (SimpleImputer method, from Scikit-learn).
183 Moreover, we used SMOTE (v0.12.0)²⁴ to oversample the minority class (PAD), to match
184 the majority class (No PAD). For the NHANES dataset, we used the 10-fold cross-validation
185 strategy, and for each fold, we filled the missing values and oversampled the minority class
186 using the same methods, and performed the so-called one-hot encoding, where we encoded
187 the categorical variables into numerical values. Next, for each dataset separately, we trained
188 the XGBoost classifier (v2.0.3)²⁵ and predicted the class of the instances on the test set,
189 from where we derived the balanced accuracy, the F1 Score, and quantified the feature
190 importance. Finally, we also implemented a version of the ML framework that allowed us to
191 assess the variable importance using SHAP (SHapley Additive exPlanations, v0.44.0).²⁶
192 Here, we performed the same steps for both HEIST and NHANES datasets, except that we
193 did not use any split strategy; instead, we filled the missing values, oversampled the minority
194 class and trained the classifier using the complete dataset. The ML pipeline is shown in
195 Figure 2A.

196

197 **Regression analysis**

198 For the regression analysis, we first tested the interaction effect of each of the iron
199 biomarkers with sex in both cohorts as iron metabolism is hallmarked by sex-specific
200 differences. Further, in the NHANES cohort, we conducted multivariable regression analyses
201 by flexibly modelling the probability of PAD's presence with one of the iron biomarkers (iron,
202 ferritin, or TSAT) by spline regression (degree of freedom=4). We chose variables from the
203 feature importance in ML analysis as covariates in the regression models on the NHANES
204 data: age, CRP, HDLc, triacylglycerols, HbA1c, ACR, BMI, ethnicity (non-Hispanic whites as
205 the reference), smoking (non-smokers as reference), the presence of diabetes, AST, ALT,
206 Hb, creatinine, antihyperlipidemic medications; the former seven variables as continuous
207 and the rest as categorical). All continuous variables were mean-centered for the analysis,
208 and variables were included in the models only when their missing data was <10%. In
209 addition, sample weights were included in the analyses of NHANES data.

210

211 All regression analyses were performed using R Statistical Software (v4.2.2²⁷). We used
212 Python (v3.10.12²⁸), and Scikit-learn (v1.4.0²⁹) for the implementation of the ML framework.
213 The reporting of the study is as per The Strengthening the Reporting of Observational

214 Studies in Epidemiology (STROBE) guidelines³⁰. We have also developed a data exploration
215 tool for readers to explore the datasets and interactively create their visualizations from the
216 paper (https://anandr.shinyapps.io/NHANES_HEISTDiC/) using Quarto³¹ and Shiny³².

217

218 **Results**

219 *Demographics*

220 The baseline characteristics of the HEIST-DiC participants are shown in Table 1. The overall
221 prevalence of PAD was 6.6% with a higher proportion among males (17/192) than females
222 (6/165). Among both sexes, the participants with PAD were older and diabetes and
223 hypertension was a common feature (Table 1). The sICAM1 and sVCAM1 levels were
224 significantly higher only among males with PAD (Table 1). Among the iron biomarkers,
225 ferritin was elevated among females, but not males, with PAD (Median 141, IQR 93-254,
226 $p=0.035$) while iron, transferrin, and TSAT were similar between HEIST-DiC participants with
227 and without PAD. (Table 1).

228

229 The analysis of NHANES builds on a previous report³³ by including data from the NHANES
230 2003-2004 cycle. The baseline characteristics of the NHANES participants included in the
231 study are shown in Table 2. Overall, the prevalence of PAD in the population was 6.25%
232 with a significantly higher prevalence among females (males: 4.96%, females: 7.9%;
233 $p=0.001$). Individuals with PAD showed differences in ethnicity, sex, presence of diabetes,
234 smoking, alcohol status, physical activity, and use of antihyperlipidemic and antiplatelet
235 medications (Table 2). Among the iron biomarkers, serum iron (Median 80, IQR 64-108
236 $\mu\text{g/dL}$; $p<0.001$) and transferrin saturation (Median 24, IQR 17-30; $p=0.021$) were reduced in
237 males with PAD while females with PAD showed higher ferritin (Median 87, IQR 56-252
238 ng/mL ; $p<0.001$) and lower TIBC (Median 60, IQR 54-68; $p<0.001$). Differences were also
239 observed in some laboratory parameters within both sexes with PAD (e.g. albumin, AST,
240 urea, creatinine, CRP, and ACR; Table 2).

241

242 Establishing a machine learning predictive framework for PAD

243 We next aimed to establish a predictive ML framework to anticipate the occurrence of PAD
244 based on biomarkers. A so-called ML classifier is trained on a dataset containing instances
245 with known outcomes (i.e., the presence or absence of PAD), and its relationship with input
246 features. During the test phase, the trained classifier is used to predict PAD in participants
247 unseen during the training phase, where we purposefully hid the true diagnostic. The result
248 is then compared to the real diagnosis, and the quality of the framework is evaluated.
249 Overall, this process enables the identification of subtle correlations between variables that

might be imperceptible to human analysis, but that can be used to anticipate PAD and enhance our diagnostic capabilities.

Our framework received the available molecular and physiological biomarkers from participants from two distinct cohorts, HEIST-DiC and NHANES as input. These cohorts have unique characteristics that require the development of separate data preprocessing, training, and testing regimens. First, the HEIST-DiC cohort comprised 357 instances (23 PAD and 334 No PAD cases) and had 35 numerical and categorical features. Here, for the training and testing of the ML framework, we used a leave-one-out split methodology. This approach involves using each instance as a test case once, while the remainder forms the training set, thus maximizing the use of available data for training. In contrast, the NHANES cohort was considerably larger, with 1188 instances (100 PAD and 1088 No PAD), and 28 features; hence, we used the 10-fold cross-validation strategy. This method divides the dataset into ten parts, with each part successively used for testing while the other 9 parts are used for training. Together, these strategies ensured robustness and reliability in our model's predictive performance.

We found that the ML framework achieved a balanced accuracy of 95% for the HEIST cohort (Figure 2B-D) and 96% for the NHANES cohort (Figure 2E-G). Balanced accuracy is an important metric in the context of imbalanced datasets, where one of the classes has many times more instances than the other (in the case of HEIST-DiC, No PAD cases were 14 times more frequent, and in NHANES, 10 times). Thus, the balanced accuracy represents the average accuracy obtained on both classes (PAD and No PAD), which in turn ensures that our model is equally adept at identifying both conditions.

Moreover, the ML framework also achieved high F1 scores of 0.89 for HEIST-DiC and 0.93 for NHANES, pointing to a strong balance between precision (the proportion of true positive results in all positive predictions), and recall (the proportion of true positive results in all actual positives). Thus, the F1 score in combination with the balanced accuracy provides realistic measures of a model's strength, particularly in datasets where one class might significantly outnumber the other. In summary, the ML framework was effective in predicting the occurrence of PAD across different cohorts, even in the face of data imbalance. Further, sex and iron biomarkers were ranked among the top predictive markers for the presence of PAD in the HEIST-DiC (ferritin, iron and transferrin) and the NHANES cohorts (ferritin, TIBC, TSAT and iron; Supplementary Figure 2; Figure 3A-B).

Regression analyses

We next performed regression analyses on the HEIST-DiC data and tested the interaction effect of each of the iron biomarkers with sex and found that ferritin showed a positive association in females with PAD (Ferritin*Sex, OR 2.68, 95% CI 0.94-7.61, $p=0.057$) while, iron (Iron*Sex, OR 1.17, 95% CI 0.44-2.96, $p=0.7$), transferrin (Transferrin*Sex, OR 1.28, 95% CI 0.48-3.17, $p=0.6$) and TSAT (TSAT*Sex, OR 1.1, 95% CI 0.35-3.2, $p=0.9$) were non-significant. Consistent with our observations in the HEIST-DiC cohort, we found that ferritin also showed a significant interaction with sex in the NHANES cohort (Ferritin*Sex, OR 1.76, 95% CI 1.16-2.67, $p=0.008$). TIBC also had a significant negative interaction (TIBC*Sex, OR 0.61, 95% CI 0.4-0.92, $p=0.018$), while iron (Iron*Sex, OR 0.95, 95% CI 0.69-1.31, $p=0.76$) and TSAT (TSAT*Sex, OR 1.15, 95% CI 0.84-1.58, $p=0.37$) were nonsignificant.

In the multivariable regression analyses of the NHANES data, we found that, ferritin or sex alone do not show a statistically significant relationship with PAD. On the other hand, the interaction between ferritin and sex reveals a significant positive association for two spline components (Table 3). Specifically, ferritin showed positive and statistically significant effects in females with two positions (48-97 ng/mL: OR 14.59, 95% CI 1.6-135.93, $p=0.019$; 98-169 ng/mL: OR 171.07, 95% CI 1.27-23404, $p=0.039$); this suggests that the relationship between PAD and ferritin is nonlinear and that within high ferritin ranges females show a positive association with PAD. Other significant covariates include age, CRP, and smoking, all of which also positively impact the outcome. By contrast, no associations were significant for iron and TSAT (Supplementary Table 2).

Discussion

Our study provides a comparison between ML and regression analyses to demonstrate the association between iron biomarkers in the context of PAD. Taken together, our findings from both analyses show that serum iron biomarkers are associated with PAD, with ferritin demonstrating a consistent association specifically in females with the two methodologies applied.

The positive association for higher serum ferritin among females with PAD strengthens previous observations on the NHANES 1999-2001 cohort³³. When this observation is compared to other related disease contexts (e.g. carotid/coronary atherosclerosis/cardiovascular disease), ferritin's positive association agrees with some studies^{11,15,16}, but not others³⁴⁻³⁶. Our findings also agree with The Iron and Atherosclerosis Study (FeAST; <https://clinicaltrials.gov>, NCT00032357), a major trial that investigated the benefits of iron reduction on the clinical outcomes in a cohort of individuals with PAD³⁷. While the initial analysis of the FeAST trial did not report a favorable outcome following

323 phlebotomy³⁷, the reanalysis of the results adjusting for the effects of age³⁸ and smoking³⁹
 324 have suggested that iron reduction (with ferritin and TSAT as endpoints) could be of
 325 significant benefit for younger individuals and smokers. However, the FeAST trial did not
 326 contain sufficient participants to study sex-specific differences. Our study also advances
 327 current knowledge. We extend the analyses by Menke et al³³, by including additional
 328 participants from the NHANES 2003-2004 cycle in the analysis. In contrast to the FeAST
 329 trial³⁷, we have included participants with prediabetes, adjusted more stringently for
 330 inflammation [excluded samples with features of inflammation (ferritin>400 ng/mL or CRP>3
 331 mg/dL) and used CRP as a covariate] and increased statistical power to study sex-specific
 332 differences.

333

334 If the high ferritin values observed in our study were solely interpreted as an indicator of
 335 body iron stores, the data suggest that there is an increased risk of PAD at the upper end of,
 336 what is clinically accepted as “physiological range”. However, ferritin is also an acute-phase
 337 protein and inflammation influences most iron parameters. During inflammatory states, iron
 338 export via ferroportin from macrophages and duodenal enterocytes is reduced by hepcidin-
 339 controlled⁴⁰ and transcriptional mechanisms^{41,42}. Consequently, macrophages become iron
 340 loaded contributing to elevated circulating ferritin levels and hypoferremia (low iron levels in
 341 the plasma)⁴³. Therefore, an alternative interpretation of our findings is that elevated ferritin
 342 may be an epiphenomenon secondary to inflammation. In addition, the presence of
 343 microinflammation in tissues cannot be completely excluded from the available data.

344

345 To summarize, our study adds to the body of evidence suggesting sex-specific differences in
 346 PAD⁴⁴ and highlights a possible role of iron (directly or indirectly) in this relationship. We
 347 have partly controlled for hormonal influences by excluding NHANES participants who are
 348 currently pregnant, menstruating, lactating, or on hormonal contraceptives, but the hormone
 349 levels were not available to be included as covariates. It is further possible that additional
 350 sex-specific mechanisms are in play.

351 In both cohorts analyzed, we observed a similar prevalence of PAD (~6%) associated with
 352 well-known risk factors (old age, diabetes, and hypertension). Markers related to metabolic
 353 and vascular health, such as blood pressure, plasma glucose levels and CRP, were elevated
 354 among PAD participants in both cohorts. Additionally, BMI was higher in PAD participants
 355 compared to non-PAD participants in HEIST-DiC, while NHANES shows more consistent
 356 BMI values across groups, with a minor increase only in females with PAD. It is important to
 357 note that the cohorts analyzed differed in that the HEIST-DiC cohort consisted of mainly
 358 European individuals with diabetes and prediabetes from a hospital-based study, while

359 NHANES is a survey cohort, representative of the US population, and therefore contains a
360 mix of ethnicities.

361

362 The strength of our study lies in the analyses of these two independent cohorts by
363 orthogonal approaches, allowing for a generalization of the findings, an increased
364 robustness of the data, and a reduction of cohort-specific biases. The examination of
365 multiple iron biomarkers, including serum iron, ferritin, TSAT, and TIBC provides a
366 comprehensive analysis of the relationship between iron and clinically apparent PAD. To the
367 best of our knowledge, this is the first study to evaluate the relationship between iron
368 biomarkers and PAD in prediabetes and diabetes. Our approach also underscores the
369 complementary nature of ML and regression analyses: while ML excels at predictive
370 accuracy and feature selection, regression analyses offer insights into interaction effects and
371 the nuanced relationship between predictors and outcomes. Therefore, integrating both
372 methodologies can leverage the strengths of each approach to provide a comprehensive
373 understanding of complex disease dynamics. Limitations include the cross-sectional nature
374 of the study which limits causal inferences and precludes the assessment of temporal
375 relationships. Further, we were unable to control for all possible confounding factors, such as
376 genetic predisposition, lifestyle, medications, or other comorbidities. The study described
377 here also did not investigate potential mechanisms underlying the observed associations.

378

379 **Conclusion**

380 Overall, our data show that elevated ferritin within the physiological range is associated with
381 clinically apparent PAD in females but not males. Further studies are necessary to confirm
382 whether the observed associations are indeed causal and to explore the potential
383 mechanisms underlying this relationship. Our findings need to be considered explorative to
384 encourage the investigation of potential biological pathways underlying the sex-specific
385 differences in iron's influence on atherosclerosis.

386

387 **Ethics statement**

388

389 The ethics committees of the University of Heidelberg approved this study (Decision No.
390 204/2004, 400/2010, and S-383/2016) according to the Declaration of Helsinki. All
391 participants entered the study according to the guidelines of the local ethics committees
392 following written informed consent to participate.

393

394 **Funding**

395

The HEIST-DiC study was funded by the Deutsche Forschungsgemeinschaft (SFB1118). MUM acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) (SFB1118; GRK 2727; FerrOs - FOR5146; Ferroptosis SPP2306: Project No.461704553) and from the Federal Ministry of Education and Research (NephRESA Nr 031L0191C; Translational Lung Research Center Heidelberg (TLRC-H), German Center for Lung Research (DZL); Project No. FKZ 82DZL004A1).

402

403

404 **Author contributions**

405

406 Anand Ruban Agarvas: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing- Original draft, Visualization

408 Stefan Kopf: Investigation, Resources, Writing - Review & Editing

409 Tiago J.S Lopes: Formal analysis, Investigation, Writing- Original draft, Visualization

410 Paul Thalmann: Validation, Writing - Review & Editing

411 José Manuel Fernández-Real: Methodology, Writing - Review & Editing, Supervision

412 Peter Nawroth: Project administration, Funding acquisition, Supervision, Writing - Review & Editing

414 Martina U. Muckenthaler: Conceptualization, Methodology, Supervision, Writing - Review & Editing, Project administration, Funding acquisition

416

417

418 **Conflict of interest**

419 None to declare

420

421 **Data and Code Availability Statement**

422 The data from this study can be explored here:

423 https://anandr.shinyapps.io/NHANES_HEISTDiC/. The HEIST-DiC dataset analyzed in the current study is available from the corresponding authors on reasonable request. The NHANES dataset is publicly available. Codes used for data analysis and visualization are accessible here (<https://github.com/griffindoc/pad>)

427

428 **Funding**

429

430 **References**

431

- 432 1. Aday AW, Matsushita K. Epidemiology of Peripheral Artery Disease and Polyvascular Disease. *Circ Res*. 2021;128(12):1818-1832. doi:10.1161/CIRCRESAHA.121.318535

433

- 434 2. Fowkes FGR, Rudan D, Rudan I, Aboyans V, Denenberg JO, McDermott MM, Norman PE,
435 Sampson UK, Williams LJ, Mensah GA, Criqui MH. Comparison of global estimates of
436 prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review
437 and analysis. *The Lancet*. 2013;382(9901):1329-1340. doi:10.1016/S0140-6736(13)61249-0
- 438 3. Galy B, Conrad M, Muckenthaler M. Mechanisms controlling cellular and systemic iron
439 homeostasis. *Nat Rev Mol Cell Biol*. Published online October 2, 2023. doi:10.1038/s41580-
440 023-00648-1
- 441 4. Sullivan JeromeL. IRON AND THE SEX DIFFERENCE IN HEART DISEASE RISK. *The Lancet*.
442 1981;317(8233):1293-1294. doi:10.1016/S0140-6736(81)92463-6
- 443 5. Bai T, Li M, Liu Y, Qiao Z, Wang Z. Inhibition of ferroptosis alleviates atherosclerosis through
444 attenuating lipid peroxidation and endothelial dysfunction in mouse aortic endothelial cell.
445 *Free Radic Biol Med*. 2020;160:92-102. doi:10.1016/j.freeradbiomed.2020.07.026
- 446 6. Moreau KL, Stauffer BL, Kohrt WM, Seals DR. Essential Role of Estrogen for Improvements in
447 Vascular Endothelial Function With Endurance Exercise in Postmenopausal Women. *The*
448 *Journal of Clinical Endocrinology & Metabolism*. 2013;98(11):4507-4515. doi:10.1210/jc.2013-
449 2183
- 450 7. Taddei S, Virdis A, Ghiadoni L, Mattei P, Sudano I, Bernini G, Pinto S, Salvetti A. Menopause Is
451 Associated With Endothelial Dysfunction in Women. *Hypertension*. 1996;28(4):576-582.
452 doi:10.1161/01.HYP.28.4.576
- 453 8. Yang Q, Jian J, Katz S, Abramson SB, Huang X. 17 β -Estradiol Inhibits Iron Hormone Hhepcidin
454 Through an Estrogen Responsive Element Half-Site. *Endocrinology*. 2012;153(7):3170-3178.
455 doi:10.1210/en.2011-2045
- 456 9. Li X, Rhee DK, Malhotra R, Mayeur C, Hurst LA, Ager E, Shelton G, Kramer Y, McCulloh D, Keefe
457 D, Bloch KD, Bloch DB, Peterson RT. Progesterone receptor membrane component-1 regulates
458 hepcidin biosynthesis. *Journal of Clinical Investigation*. 2015;126(1):389-401.
459 doi:10.1172/JCI83831
- 460 10. Vinchi F, Porto G, Simmelbauer A, Altamura S, Passos ST, Garbowski M, Silva AMN, Spaich S,
461 Seide SE, Sparla R, Hentze MW, Muckenthaler MU. Atherosclerosis is aggravated by iron
462 overload and ameliorated by dietary and pharmacological iron restriction. *Eur Heart J*.
463 Published online March 20, 2019. doi:10.1093/eurheartj/ehz112
- 464 11. Wolff B, Völzke H, Lüdemann J, Robinson D, Vogelgesang D, Staudt A, Kessler C, Dahm JB, John
465 U, Felix SB. Association between high serum ferritin levels and carotid atherosclerosis in the
466 study of health in Pomerania (SHIP). *Stroke*. 2004;35(2):453-457.
467 doi:10.1161/01.STR.0000114875.31599.1C
- 468 12. Sung KC, Kang SM, Cho EJ, Park JB, Wild SH, Byrne CD. Ferritin Is Independently Associated
469 With the Presence of Coronary Artery Calcium in 12 033 Men. *ATVB*. 2012;32(10):2525-2530.
470 doi:10.1161/ATVBAHA.112.253088
- 471 13. Kiechl S, Willeit J, Egger G, Poewe W, Oberhollenzer F. Body iron stores and the risk of carotid
472 atherosclerosis: prospective results from the Bruneck study. *Circulation*. 1997;96(10):3300-
473 3307. doi:10.1161/01.cir.96.10.3300

- 474 14. Das De S, Krishna S, Jethwa A. Iron status and its association with coronary heart disease:
475 systematic review and meta-analysis of prospective studies. *Atherosclerosis*. 2015;238(2):296-
476 303. doi:10.1016/j.atherosclerosis.2014.12.018
- 477 15. Guo S, Mao X, Li X, Ouyang H. Association between iron status and incident coronary artery
478 disease: a population based-cohort study. *Sci Rep*. 2022;12(1):17490. doi:10.1038/s41598-022-
479 22275-0
- 480 16. Klip IJT, Voors AA, Swinkels DW, Bakker SJL, Kootstra-Ros JE, Lam CS, Van Der Harst P, Van
481 Veldhuisen DJ, Van Der Meer P. Serum ferritin and risk for new-onset heart failure and
482 cardiovascular events in the community: Serum ferritin and risk for new-onset HF and CV
483 events. *Eur J Heart Fail*. 2017;19(3):348-356. doi:10.1002/ehf.622
- 484 17. Gill D, Brewer CF, Monori G, Trégouët DA, Franceschini N, Giambartolomei C, INVENT
485 Consortium, Tzoulaki I, Dehghan A. Effects of Genetically Determined Iron Status on Risk of
486 Venous Thromboembolism and Carotid Atherosclerotic Disease: A Mendelian Randomization
487 Study. *J Am Heart Assoc*. 2019;8(15):e012994. doi:10.1161/JAHA.119.012994
- 488 18. Altamura S, Kopf S, Schmidt J, Müdder K, da Silva AR, Nawroth P, Muckenthaler MU.
489 Uncoupled iron homeostasis in type 2 diabetes mellitus. *J Mol Med*. 2017;95(12):1387-1398.
490 doi:10.1007/s00109-017-1596-3
- 491 19. Zoller H, Schaefer B, Vanclooster A, Griffiths B, Bardou-Jacquet E, Corradini E, Porto G, Ryan J,
492 Cornberg M. EASL Clinical Practice Guidelines on haemochromatosis. *Journal of Hepatology*.
493 2022;77(2):479-502. doi:10.1016/j.jhep.2022.03.033
- 494 20. Qiu Z, Chen X, Geng T, Wan Z, Lu Q, Li L, Zhu K, Zhang X, Liu Y, Lin X, Chen L, Shan Z, Liu L, Pan
495 A, Liu G. Associations of Serum Carotenoids With Risk of Cardiovascular Mortality Among
496 Individuals With Type 2 Diabetes: Results From NHANES. *Diabetes Care*. 2022;45(6):1453-1461.
497 doi:10.2337/dc21-2371
- 498 21. Newsome PN, Cramb R, Davison SM, Dillon JF, Foulerton M, Godfrey EM, Hall R, Harrower U,
499 Hudson M, Langford A, Mackie A, Mitchell-Thain R, Sennett K, Sheron NC, Verne J, Walmsley
500 M, Yeoman A. Guidelines on the management of abnormal liver blood tests. *Gut*. 2018;67(1):6-
501 19. doi:10.1136/gutjnl-2017-314924
- 502 22. Ahlqvist E, Prasad RB, Groop L. Subtypes of Type 2 Diabetes Determined From Clinical
503 Parameters. *Diabetes*. 2020;69(10):2086-2093. doi:10.2337/dbi20-0001
- 504 23. Aboyans V, Ricco JB, Bartelink MLEL, Björck M, Brodmann M, Cohnert T, Collet JP, Czerny M, De
505 Carlo M, Debus S, Espinola-Klein C, Kahan T, Kownator S, Mazzolai L, Naylor AR, Roffi M, Röther
506 J, Sprynger M, Tendera M, Tepe G, Venermo M, Vlachopoulos C, Desormais I, ESC Scientific
507 Document Group, Widimsky P, Kolh P, Agewall S, Bueno H, Coca A, De Borst GJ, Delgado V, Dick
508 F, Erol C, Ferrini M, Kakkos S, Katus HA, Knuuti J, Lindholt J, Mattle H, Pieniazek P, Piepoli MF,
509 Scheinert D, Sievert H, Simpson I, Sulzenko J, Tamargo J, Tokgozoglu L, Torbicki A, Tsakountakis
510 N, Tuñón J, De Céniga MV, Windecker S, Zamorano JL, Windecker S, Aboyans V, Agewall S,
511 Barbato E, Bueno H, Coca A, Collet JP, Coman IM, Dean V, Delgado V, Fitzsimons D, Gaemperli
512 O, Hindricks G, Iung B, Juni P, Katus HA, Knuuti J, Lancellotti P, Leclercq C, McDonagh T, Piepoli
513 MF, Ponikowski P, Richter DJ, Roffi M, Shlyakhto E, Simpson IA, Zamorano JL, Zelveian PH,
514 Haumer M, Isachkin D, De Backer T, Dilic M, Petrov I, Kirhmajer MV, Karetova D, Prescott E,
515 Soliman H, Paapstel A, Makinen K, Tosev S, Messas E, Pagava Z, Müller OJ, Naka KK, Járαι Z, et
516 al. 2017 ESC Guidelines on the Diagnosis and Treatment of Peripheral Arterial Diseases, in

- 517 collaboration with the European Society for Vascular Surgery (ESVS). *European Heart Journal*.
518 2018;39(9):763-816. doi:10.1093/eurheartj/ehx095
- 519 24. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling
520 Technique. *Jair*. 2002;16:321-357. doi:10.1613/jair.953
- 521 25. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd*
522 *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM;
523 2016:785-794. doi:10.1145/2939672.2939785
- 524 26. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. Published online
525 2017. doi:10.48550/ARXIV.1705.07874
- 526 27. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2022.
527 <https://www.R-project.org/>
- 528 28. Rossum G van, Drake FL. *The Python Language Reference*. Release 3.0.1 [Repr.]. Python
529 Software Foundation; 2010.
- 530 29. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer
531 P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay
532 É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*.
533 2011;12(85):2825-2830.
- 534 30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative.
535 The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)
536 statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453-1457.
537 doi:10.1016/S0140-6736(07)61602-X
- 538 31. Allaire J. quarto: R Interface to “Quarto” Markdown Publishing System_. R package version 1.2.
539 Published online 2022. <https://CRAN.R-project.org/package=quarto>
- 540 32. Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A, Borges
541 B. Shiny: Web Application Framework for R_. R package version 1.7.4.1. Published online 2023.
542 <https://CRAN.R-project.org/package=shiny>
- 543 33. Menke A, Fernández-Real JM, Muntner P, Guallar E. The association of biomarkers of iron
544 status with peripheral arterial disease in US adults. *BMC Cardiovasc Disord*. 2009;9(1):34.
545 doi:10.1186/1471-2261-9-34
- 546 34. Zhou J, Zhao R, Wang D, Gao Q, Zhao D, Ouyang B, Hao L, Peng X. Sex-Specific Association
547 Between Iron Status and the Predicted 10-Year Risk for Atherosclerotic Cardiovascular Disease
548 in Hypertensive Patients. *Biol Trace Elem Res*. 2022;200(11):4594-4607. doi:10.1007/s12011-
549 021-03060-y
- 550 35. Ellervik C, Marott JL, Tybjaerg-Hansen A, Schnohr P, Nordestgaard BG. Total and cause-specific
551 mortality by moderately and markedly increased ferritin concentrations: general population
552 study and metaanalysis. *Clin Chem*. 2014;60(11):1419-1428.
553 doi:10.1373/clinchem.2014.229013
- 554 36. Kadoglou NPE, Biddulph JP, Rafnsson SB, Trivella M, Nihoyannopoulos P, Demakakos P. The
555 association of ferritin with cardiovascular and all-cause mortality in community-dwellers: The

556 English longitudinal study of ageing. *PLoS One*. 2017;12(6):e0178994.
557 doi:10.1371/journal.pone.0178994

558 37. Zacharski LR, Chow BK, Howes PS, Shamayeva G, Baron JA, Dalman RL, Malenka DJ, Ozaki CK,
559 Lavori PW. Reduction of Iron Stores and Cardiovascular Outcomes in Patients With Peripheral
560 Arterial Disease: A Randomized Controlled Trial. *JAMA*. 2007;297(6):603.
561 doi:10.1001/jama.297.6.603

562 38. Zacharski LR, Shamayeva G, Chow BK. Effect of controlled reduction of body iron stores on
563 clinical outcomes in peripheral arterial disease. *American Heart Journal*. 2011;162(5):949-
564 957.e1. doi:10.1016/j.ahj.2011.08.013

565 39. DePalma RG, Zacharski LR, Chow BK, Shamayeva G, Hayes VW. Reduction of iron stores and
566 clinical outcomes in peripheral arterial disease: outcome comparisons in smokers and non-
567 smokers. *Vascular*. 2013;21(4):233-241. doi:10.1177/1708538113478776

568 40. Nemeth E, Tuttle MS, Powelson J, Vaughn MB, Donovan A, Ward DM, Ganz T, Kaplan J.
569 Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization.
570 *Science*. 2004;306(5704):2090-2093. doi:10.1126/science.1104742

571 41. Guida C, Altamura S, Klein FA, Galy B, Boutros M, Ulmer AJ, Hentze MW, Muckenthaler MU. A
572 novel inflammatory pathway mediating rapid hepcidin-independent hypoferremia. *Blood*.
573 2015;125(14):2265-2275. doi:10.1182/blood-2014-08-595256

574 42. Marques O, Weiss G, Muckenthaler MU. The role of iron in chronic inflammatory diseases:
575 from mechanisms to treatment options in anemia of inflammation. *Blood*. 2022;140(19):2011-
576 2023. doi:10.1182/blood.2021013472

577 43. Cohen LA, Gutierrez L, Weiss A, Leichtmann-Bardoogo Y, Zhang D liang, Crooks DR, Sougrat R,
578 Morgenstern A, Galy B, Hentze MW, Lazaro FJ, Rouault TA, Meyron-Holtz EG. Serum ferritin is
579 derived primarily from macrophages through a nonclassical secretory pathway. *Blood*.
580 2010;116(9):1574-1584. doi:10.1182/blood-2009-11-253815

581 44. Pabon M, Cheng S, Altin SE, Sethi SS, Nelson MD, Moreau KL, Hamburg N, Hess CN. Sex
582 Differences in Peripheral Artery Disease. *Circ Res*. 2022;130(4):496-511.
583 doi:10.1161/CIRCRESAHA.121.320702

584

Legends

Table 1. Demographics of HEIST-DiC participants

The demographics of the HEIST-DiC participants classified by sex and the presence of PAD are shown in Table 1. For continuous variables, the summary data are shown as Mean±Standard Deviation and for categorical variables as N (%). P-values as calculated by the Wilcoxon rank sum test or Pearson's Chi-squared test^C and significant P-values are boldfaced.

BMI: Body Mass Index, HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, HbA1c: Glycated hemoglobin, ICAM1: Intercellular Adhesion Molecule 1, VCAM1: Vascular Cell Adhesion Molecule 1, hsCRP: High sensitivity C-reactive protein, IL6: Interleukin 6, hsTNT: High sensitivity Troponin T, proBNP: N-Terminal B-type Natriuretic Peptide, eGFR: Estimated Glomerular Filtration Rate, TSAT: Transferrin Saturation

Table 2. Demographics of NHANES 1999-2004 participants

The demographics of the NHANES 1999-2004 participants classified by sex and the presence of PAD are shown in Table 2. The ¹N of participants shown are unweighted numbers. For continuous variables, the summary data are shown as Median (Interquartile Range), and for categorical variables as N (%)- survey weights were included in calculating the summary data. P-values as calculated by the Wilcoxon rank-sum test for complex survey samples or chi-squared test with Rao & Scott's second-order correction^C and significant P-values are boldfaced.

BMI: Body Mass Index, HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, HbA1c: Glycated hemoglobin, CRP: High sensitivity C-reactive protein, eGFR: Estimated Glomerular Filtration Rate, TSAT: Transferrin Saturation, TIBC: Total Iron Binding Capacity, AST: Aspartate Aminotransferase, ALT: Alanine Aminotransferase, FIB4 Score: Fibrosis-4 index Score

Table 3. Multivariable regression analysis of ferritin and PAD in the NHANES Cohort

We conducted multivariable regression analyses by flexibly modelling the probability of PAD's presence with ferritin by spline regression (degree of freedom=4). The following variables were included a covariates: age, CRP, HDLc, triacylglycerols, HbA1c, ACR, BMI, ethnicity (non-Hispanic whites as the reference), smoking (non-smokers as reference), the presence of diabetes, AST, ALT, Hb, creatinine, antihyperlipidemic medications; the former seven variables as continuous and the rest as categorical). All continuous variables were mean-centered for the analysis. In addition, sample weights were included in the analyses BMI: Body Mass Index, HDL: High-Density Lipoprotein, CRP: High sensitivity C-reactive protein, AST: Aspartate Aminotransferase, ALT: Alanine Aminotransferase; ACR: Albumin-creatinine ratio, OR: Odds Ratio, 95% CI: 95% Confidence Interval, SE: Standard Error

Figure 1. Outline of the study and analysis approach

Figure 2. Setup and results of the machine learning approach

(A) The datasets were divided using either cross-fold validation (NHANES), or the leave-one-out strategy (HEIST-DiC). Next, the categorical variables were transformed into binary values (the so-called 'hot-encode'), we filled missing values and oversampled the minority classes (see Methods). After training the classifier, the instances of the test set had their classes predicted, and we evaluated the results using classical ML metrics. (B-D) The ROC and precision-recall curves of the HEIST-DiC cohort data, and its confusion matrix. (E-G) The ROC and precision-recall curves of the NHANES cohort data, and its confusion matrix.

Figure 3. Feature importance of machine learning classification

(A-B) We used the SHAP method to quantify the importance of the features that were contributing the most to the prediction of PAD occurrence. Here, higher values indicate a

higher contribution of a given feature. In both panels, iron-related markers are color-coded, and appear among the top features in terms of relevance to the prediction of PAD.

Supplementary files

Supplementary Table 1. Data fields and variables from NHANES 1999-2004.

The table shows the variables and the corresponding codes used for accessing data from the NHANES. The hyperlinks to the procedure manuals and datafiles are also provided in the table.

Supplementary Table 2. Multivariable regression analysis of iron biomarkers and PAD in the NHANES Cohort

We conducted multivariable regression analyses by flexibly modelling the probability of PAD's presence with iron biomarkers (iron or TSAT) by spline regression (degree of freedom=4). The following variables were included as covariates: age, CRP, HDLc, triacylglycerols, HbA1c, ACR, BMI, ethnicity (non-Hispanic whites as the reference), smoking (non-smokers as reference), the presence of diabetes, AST, ALT, Hb, creatinine, antihyperlipidemic medications; the former seven variables as continuous and the rest as categorical). All continuous variables were mean-centered for the analysis. In addition, sample weights were included in the analyses

BMI: Body Mass Index, HDL: High-Density Lipoprotein, CRP: High sensitivity C-reactive protein, AST: Aspartate Aminotransferase, ALT: Alanine Aminotransferase; ACR: Albumin-creatinine ratio, OR: Odds Ratio, 95% CI: 95% Confidence Interval, SE: Standard Error

Supplementary Figure 1. Feature importance of xGBoost classification

(A-B) We used the XGBoost classifier and its output to quantify the importance of the features to the prediction of PAD occurrence. In both panels, higher values indicate a higher contribution of a given feature, and iron-related markers are color coded, and appear among the top features in terms of relevance to the prediction of this condition.

No	129 (67%)	118 (67%)	11 (65%)	>0.9	136 (82%)	132 (83%)	4 (67%)	0.6
Yes	63 (33%)	57 (33%)	6 (35%)		29 (18%)	27 (17%)	2 (33%)	
Insulin use								
No	157 (82%)	145 (83%)	12 (71%)	0.4	141 (85%)	137 (86%)	4 (67%)	0.5
Yes	35 (18%)	30 (17%)	5 (29%)		24 (15%)	22 (14%)	2 (33%)	
Fasting blood glucose (mg/dL)	127 (103, 161)	125 (102, 154)	167 (123, 204)	0.002	106 (92, 145)	105 (92, 139)	152 (139, 174)	0.067
HbA1c (g/dL)	6.40 (5.70, 7.50)	6.40 (5.60, 7.40)	7.50 (6.30, 7.90)	0.014	5.80 (5.40, 6.80)	5.80 (5.40, 6.80)	7.10 (6.60, 8.20)	0.056
Parameters of vascular function								
Ankle-brachial index (Right)	1.10 (1.03, 1.16)	1.11 (1.06, 1.16)	0.86 (0.76, 0.88)	<0.001	1.08 (1.02, 1.15)	1.09 (1.03, 1.15)	0.82 (0.76, 0.86)	<0.001
Ankle-brachial index (Left)	1.10 (1.05, 1.16)	1.11 (1.06, 1.17)	0.92 (0.88, 0.95)	<0.001	1.09 (1.04, 1.15)	1.09 (1.05, 1.15)	0.97 (0.86, 0.98)	<0.001
Carotid intima-media thickness (mm)	0.60 (0.54, 0.70)	0.60 (0.54, 0.70)	0.70 (0.59, 0.85)	0.025	0.60 (0.50, 0.66)	0.58 (0.50, 0.65)	0.83 (0.70, 0.98)	0.002
Carotid-femoral Pulse Wave Velocity (m/s)	8.70 (7.70, 9.80)	8.70 (7.60, 9.70)	9.75 (9.55, 11.50)	0.004	8.10 (7.10, 9.70)	8.10 (7.10, 9.60)	11.35 (11.00, 12.60)	0.002
ICAM1 (ng/mL)	294 (212, 366)	289 (206, 357)	380 (303, 419)	0.008	305 (225, 363)	305 (229, 367)	225 (218, 363)	0.6
VCAM1 (ng/mL)	847 (692, 1,052)	827 (682, 1,033)	1,104 (878, 1,330)	0.017	804 (697, 961)	803 (697, 962)	856 (467, 905)	0.8

			1,189)					
Other laboratory parameters								
Triacylglycerols (mg/dL)	128 (90, 199)	125 (87, 191)	177 (148, 247)	0.011	119 (82, 166)	118 (80, 163)	177 (128, 207)	0.092
Cholesterol (mg/dL)	184 (157, 221)	184 (156, 221)	183 (159, 216)	0.8	204 (173, 231)	204 (173, 231)	184 (159, 240)	0.5
HDL (mg/dL)	46 (39, 55)	47 (39, 55)	40 (38, 47)	0.067	57 (47, 71)	58 (47, 71)	50 (37, 57)	0.3
LDL (mg/dL)	107 (82, 134)	107 (82, 135)	107 (83, 132)	0.9	116 (89, 139)	116 (90, 139)	81 (78, 116)	0.088
Hemoglobin (g/dL)	14.90 (14.10, 15.50)	14.90 (14.10, 15.50)	14.30 (13.20, 15.10)	0.14	13.60 (13.00, 14.20)	13.60 (13.00, 14.20)	13.25 (12.80, 14.00)	0.5
Mean Corpuscular Volume (fL)	89.0 (86.0, 92.0)	89.0 (86.0, 92.0)	90.0 (87.0, 92.0)	0.2	88.0 (85.0, 91.0)	88.0 (85.0, 91.0)	87.5 (84.0, 91.0)	0.7
hsCRP (mg/L)	1.09 (0.52, 2.54)	1.05 (0.51, 2.35)	1.32 (0.79, 4.36)	0.3	2.1 (0.7, 4.7)	1.9 (0.7, 4.6)	5.0 (2.9, 15.1)	0.083
IL6 (pg/mL)	2.45 (2.00, 3.70)	2.35 (2.00, 3.60)	3.20 (2.40, 4.20)	0.010	2.55 (2.00, 4.15)	2.50 (2.00, 4.10)	3.80 (2.50, 6.90)	0.3
hsTNT (ng/L)	9.0 (7.0, 13.0)	9.0 (7.0, 12.0)	11.0 (9.0, 20.0)	0.033	6.0 (4.0, 8.0)	6.0 (4.0, 8.0)	9.0 (8.0, 9.0)	0.2
proBNP (pg/mL)	60 (31, 129)	59 (29, 118)	113 (35, 251)	0.052	71 (39, 130)	69 (37, 121)	175 (135, 236)	0.008
eGFR (ml/min/1.73m²)	91 (80, 100)	92 (82, 101)	76 (62, 91)	0.002	93 (81, 101)	93 (82, 101)	64 (50, 86)	0.016

Urinary Albumin Creatinine Ratio (mg/g)	8 (4, 23)	8 (4, 18)	25 (11, 139)	0.001	7 (4, 14)	7 (4, 13)	26 (14, 111)	0.012
Iron-related parameters								
Iron (µM)	95 (75, 119)	96 (75, 120)	85 (76, 108)	0.7	90 (71, 118)	90 (70, 118)	95 (79, 126)	0.6
Ferritin (ng/mL)	155 (81, 259)	153 (78, 250)	192 (126, 306)	0.10	79 (44, 136)	77 (42, 131)	141 (90, 278)	0.031
Transferrin (g/L)	2.38 (2.18, 2.61)	2.38 (2.18, 2.61)	2.24 (2.15, 2.73)	0.6	2.42 (2.24, 2.69)	2.42 (2.24, 2.67)	2.56 (2.16, 2.88)	0.6
TSAT (%)	25 (20, 32)	26 (20, 33)	23 (21, 27)	0.2	24 (18, 31)	25 (18, 31)	22 (19, 24)	0.5

¹Median (Q1, Q3); n (%)

Table 1. Demographics of HEIST-DiC participants

The demographics of the HEIST-DiC participants classified by sex and the presence of PAD are shown in Table 1. For continuous variables, the summary data are shown as Median (Q1, Q3) and for categorical variables as N (%). P-values as calculated by the Wilcoxon rank sum test or Pearson's Chi-squared test^C and significant P-values are boldfaced.

BMI: Body Mass Index, HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, HbA1c: Glycated hemoglobin, ICAM1: Intercellular Adhesion Molecule 1, VCAM1: Vascular Cell Adhesion Molecule 1, hsCRP: High sensitivity C-reactive protein, IL6: Interleukin 6, hsTNT: High sensitivity Troponin T, proBNP: N-Terminal B-type Natriuretic Peptide, eGFR: Estimated Glomerular Filtration Rate, TSAT: Transferrin Saturation

	Males				Females			
Characteristic	Overall N = 2792 ¹	No PAD, N = 2573 ¹	1, N = 219 ¹	p- value ²	Overall N =2309 ¹	0, N = 2082 ¹	1, N = 227 ¹	p- value ²
Ethnicity								
Non-Hispanic White	1459 (78%)	1330 (78%)	129 (83%)	0.003	1089 (73%)	971 (72%)	118 (76%)	<0.001
Non-Hispanic Black	497 (9.0%)	454 (8.9%)	43 (11%)		475 (11%)	418 (11%)	57 (16%)	
Mexican- American	662 (5.6%)	619 (5.6%)	43 (5.2%)		546 (5.0%)	504 (5.1%)	42 (4.2%)	
Others	174 (7.6%)	170 (7.9%)	4 (1.2%)		199 (11%)	189 (12%)	10 (4.0%)	
Age (years)	52 (45, 61)	51 (45, 60)	68 (60, 76)	<0.001	52 (45, 65)	51 (45, 64)	70 (54, 81)	<0.001
BMI (kg/sq.m)	27.9 (25.1, 31.0)	27.9 (25.1, 30.9)	27.8 (25.0, 32.2)	>0.9	27 (24, 32)	27 (24, 32)	29 (25, 34)	0.034
Diabetes (coded)								
No diabetes	1796 (75%)	1698 (76%)	98 (49%)	<0.001	1461 (73%)	1358 (75%)	103 (51%)	<0.001
Prediabetes	444 (13%)	387 (12%)	57 (27%)		418 (16%)	357 (15%)	61 (25%)	
Diabetes	450 (12%)	395 (12%)	55 (24%)		349 (11%)	296 (10%)	53 (23%)	
Hypertension (coded)								
No	252 (14%)	233 (14%)	19 (13%)	0.9	209 (15%)	188 (15%)	21 (14%)	0.9
Yes	1557	1404	153 (87%)		1296	1129	167 (86%)	

	(86%)	(86%)			(85%)	(85%)		
Smoking (coded)								
Nonsmoker	974 (36%)	931 (37%)	43 (18%)	<0.001	1397 (58%)	1277 (58%)	120 (51%)	0.089
Former smoker	1134 (39%)	1023 (38%)	111 (49%)		532 (24%)	466 (23%)	66 (28%)	
Current smoker	680 (25%)	615 (25%)	65 (33%)		375 (18%)	335 (18%)	40 (21%)	
Alcohol drinker (coded)								
No	188 (33%)	174 (33%)	14 (39%)	0.5	552 (44%)	484 (43%)	68 (50%)	0.3
Yes	339 (67%)	317 (67%)	22 (61%)		570 (56%)	506 (57%)	64 (50%)	
Antihyperlipidemic medications (coded)								
No	2269 (91%)	2083 (91%)	186 (91%)	0.9	1984 (95%)	1788 (95%)	196 (91%)	0.037
Yes	177 (8.7%)	159 (8.7%)	18 (9.2%)		113 (5.2%)	97 (4.8%)	16 (9.3%)	
Antiplatelet medications (coded)								
No	2404 (99%)	2209 (99%)	195 (97%)	0.006	2075 (99%)	1867 (99%)	208 (99%)	0.3
Yes	42 (1.1%)	33 (1.0%)	9 (2.9%)		22 (0.8%)	18 (0.7%)	4 (1.5%)	
Ankle-brachial Index (Left)	1.16 (1.09, 1.22)	1.16 (1.10, 1.22)	0.88 (0.79, 1.01)	<0.001	1.11 (1.04, 1.18)	1.12 (1.06, 1.19)	0.88 (0.78, 0.94)	<0.001
Ankle-brachial	1.16 (1.10,	1.16 (1.11,	0.87 (0.75,	<0.001	1.10 (1.03,	1.11 (1.05,	0.86 (0.77,	<0.001

Index (Right)	1.22)	1.22)	0.97)		1.18)	1.18)	0.90)	
Albumin (g/L)	44.00 (42.00, 46.00)	44.00 (42.00, 46.00)	43.00 (41.00, 44.00)	<0.001	43.0 (41.0, 44.0)	43.0 (41.0, 45.0)	42.0 (39.0, 44.0)	0.001
AST (U/L)	26 (21, 33)	26 (21, 34)	22 (18, 27)	<0.001	19 (15, 25)	19 (15, 25)	18 (16, 23)	0.2
ALT (U/L)	24 (21, 29)	24 (21, 29)	23 (20, 28)	0.059	21 (18, 25)	21 (18, 25)	22 (19, 26)	0.15
Blood urea nitrogen (mmol/L)	5.00 (4.28, 6.10)	5.00 (3.93, 6.07)	6.07 (4.64, 7.14)	<0.001	4.60 (3.57, 5.71)	4.60 (3.57, 5.71)	5.00 (3.90, 6.78)	0.002
Creatinine (μmol/L)	88 (71, 97)	80 (71, 97)	88 (80, 106)	0.018	62 (53, 71)	62 (53, 71)	71 (62, 88)	<0.001
Total Cholesterol (mg/dL)	203 (179, 228)	204 (179, 228)	193 (171, 220)	0.008	207 (184, 235)	207 (183, 235)	211 (188, 238)	0.3
Triacylglycerols (mg/dL)	132 (89, 200)	132 (89, 200)	130 (86, 190)	>0.9	108 (73, 160)	107 (71, 156)	132 (95, 195)	<0.001
LDL (mg/dL)	124 (103, 147)	124 (103, 148)	115 (85, 136)	0.038	122 (101, 145)	122 (101, 145)	119 (97, 143)	0.8
HDL (mg/dL)	44 (37, 54)	44 (37, 54)	44 (37, 52)	0.4	55 (46, 67)	56 (46, 67)	52 (44, 61)	0.006
Fasting plasma glucose (mg/dL)	101 (94, 111)	100 (94, 110)	108 (92, 125)	0.3	95 (90, 104)	95 (90, 103)	102 (96, 115)	<0.001
Insulin (pmol/L)	57 (37, 92)	56 (37, 90)	85 (41, 129)	0.089	48 (33, 80)	47 (32, 77)	99 (46, 130)	0.007
CRP (mg/dL)	0.18 (0.08, 0.37)	0.18 (0.08, 0.37)	0.27 (0.14, 0.44)	<0.001	0.25 (0.10, 0.58)	0.25 (0.10, 0.57)	0.32 (0.15, 0.76)	0.036

Hemoglobin (g/dL)	15.30 (14.70, 16.00)	15.30 (14.70, 16.00)	15.00 (14.10, 15.80)	0.006	13.90 (13.10, 14.50)	13.90 (13.10, 14.50)	13.80 (12.90, 14.70)	>0.9
FIB4 Score	1.14 (0.89, 1.48)	1.13 (0.89, 1.47)	1.28 (1.02, 1.62)	0.002	0.85 (0.63, 1.15)	0.84 (0.62, 1.14)	1.07 (0.80, 1.37)	<0.001
Albumin-creatinine ratio (µg/mg)	0.59 (0.40, 1.06)	0.58 (0.40, 1.01)	1.07 (0.63, 3.09)	<0.001	0.81 (0.54, 1.54)	0.79 (0.54, 1.44)	1.30 (0.70, 2.67)	<0.001
Iron (µM)	16.5 (12.9, 20.8)	16.7 (12.9, 20.9)	15.4 (11.6, 20.4)	0.065	14.1 (10.6, 18.1)	14.3 (10.6, 18.1)	12.9 (10.2, 16.5)	0.025
Ferritin (ng/mL)	142 (93, 223)	142 (93, 223)	120 (81, 197)	0.3	60 (31, 109)	60 (31, 107)	85 (56, 192)	<0.001
TSAT (%)	26 (20, 33)	27 (20, 33)	25 (17, 32)	0.079	22 (16, 29)	22 (16, 29)	20 (16, 28)	0.2
TIBC (µM)	63 (57, 70)	63 (57, 69)	63 (57, 70)	0.8	64 (58, 71)	64 (58, 71)	61 (54, 70)	0.001

Table 2. Demographics of NHANES 1999-2004 participants

The demographics of the NHANES 1999-2004 participants classified by sex and the presence of PAD are shown in Table 2. The ¹N of participants shown are unweighted numbers. For continuous variables, the summary data are shown as Median (Q1, Q3), and for categorical variables as N (%)- survey weights were included in calculating the summary data. P-values as calculated by the Wilcoxon rank-sum test for complex survey samples or chi-squared test with Rao & Scott's second-order correction^C and significant P-values are boldfaced.

BMI: Body Mass Index, HDL: High-Density Lipoprotein, LDL: Low-Density Lipoprotein, HbA1c: Glycated hemoglobin, CRP: High sensitivity C-reactive protein, eGFR: Estimated Glomerular Filtration Rate, TSAT: Transferrin Saturation, TIBC: Total Iron Binding Capacity, AST: Aspartate Aminotransferase, ALT: Alanine Aminotransferase, FIB4 Score: Fibrosis-4 index Score

Table 3	OR	SE	Z score	P. Value	95% CI	
					Low	High
Intercept	0.001	3.205	-2.338	0.019	0.000	0.297
Ferritin (<48 ng/mL)	0.868	0.896	-0.158	0.875	0.150	5.031
Ferritin (48-96 ng/mL)	0.481	1.478	-0.495	0.621	0.027	8.714
Ferritin (97-168 ng/mL)	0.284	2.034	-0.619	0.536	0.005	15.306
Ferritin (>169 ng/mL)	0.777	1.061	-0.238	0.812	0.097	6.214
Sex:F (ref=M)	0.281	1.053	-1.205	0.228	0.036	2.215
Ferritin (<48 ng/mL)*Females	1.042	1.177	0.035	0.972	0.104	10.471
Ferritin (48-96 ng/mL) *Females	14.593	1.139	2.354	0.019	1.567	135.935
Ferritin (97-168 ng/mL) *Females	171.066	2.501	2.056	0.040	1.270	23034.375
Ferritin (>169 ng/mL) *Females	3.073	1.247	0.900	0.368	0.266	35.427
Former Smoker (ref= Never Smokers)	1.319	0.392	0.706	0.480	0.612	2.844
Current Smoker (ref= Never Smokers)	4.696	0.282	5.487	<0.001	2.703	8.160
Prediabetes (ref=No diabetes)	1.570	0.257	1.754	0.080	0.948	2.600
Diabetes Mellitus (ref=No diabetes)	1.337	0.403	0.720	0.472	0.607	2.945
Non-Hispanic Black (ref: Non-Hispanic White)	1.296	0.398	0.651	0.515	0.594	2.829
Mexican-American (ref: Non-Hispanic White)	1.281	0.311	0.797	0.425	0.697	2.355
Other Ethnicities (ref: Non-Hispanic White)	0.251	0.740	-1.865	0.062	0.059	1.072
CRP	1.810	0.259	2.293	0.022	1.090	3.004
Age	1.095	0.013	7.241	<0.001	1.068	1.122
Creatinine	1.009	0.006	1.448	0.148	0.997	1.021
ACR	0.058	1.987	-1.430	0.153	0.001	2.868
Triacylglycerols	0.997	0.068	-0.039	0.969	0.873	1.140
BMI	1.078	0.162	0.463	0.644	0.785	1.480
HDL	0.751	0.345	-0.831	0.406	0.382	1.476
Hemoglobin	0.897	0.190	-0.574	0.566	0.618	1.301
Antihyperlipidemic medications (ref=No)	1.318	0.546	0.505	0.613	0.452	3.841
AST	1.005	0.013	0.438	0.662	0.981	1.030
ALT	0.988	0.018	-0.647	0.518	0.953	1.024



HEIST-DiC Cohort



NHANES 1999-2004

Iron biomarkers

- *Iron*
- *Ferritin*
- *TSAT*
- *Transferrin*

Vascular function markers

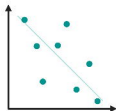
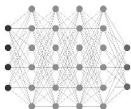
- *ICAM1*
- *VCAM1*
- *Ankle-Brachial Index*

Iron biomarkers

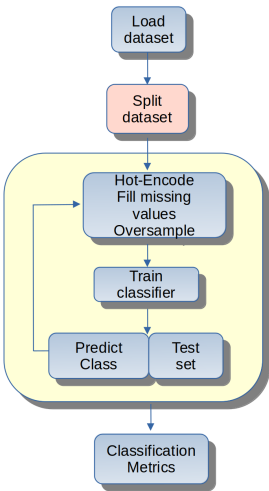
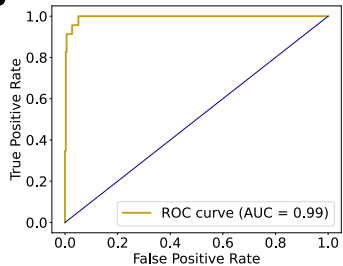
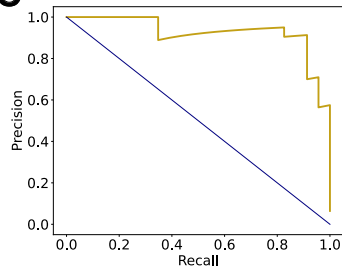
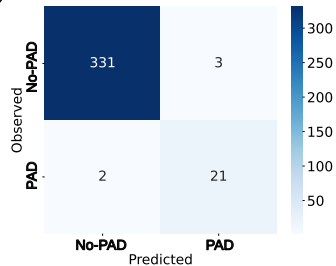
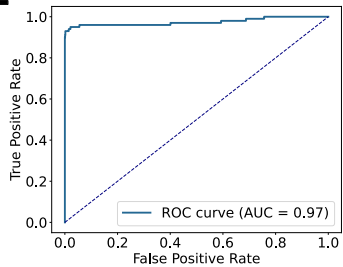
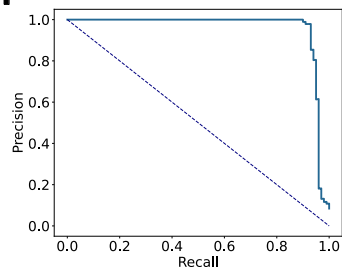
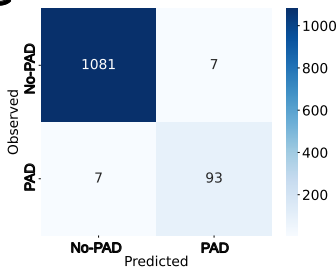
- *Iron*
- *Ferritin*
- *TSAT*
- *TIBC*

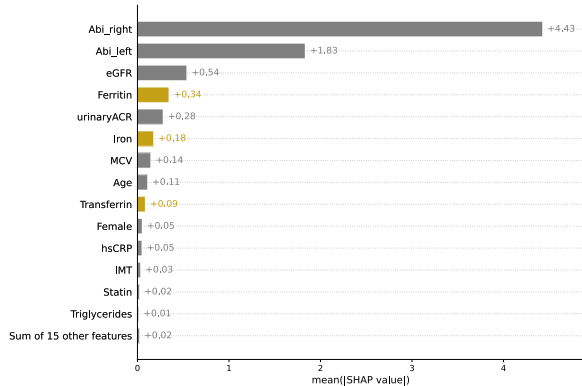
Vascular function markers

- *Ankle-Brachial Index*



Machine learning & Regression analysis

A**B****C****D****E****F****G**

A**B**