

Original research

Serum ferritin levels can predict long-term outcomes in patients with metabolic dysfunction-associated steatotic liver disease

Angelo Armandi ^{1,2}, Tiziana Sanavia ³, Ramy Younes,⁴ Gian Paolo Caviglia,¹ Chiara Rosso,¹ Olivier Govaere,⁵ Antonio Liguori,^{6,7,8} Paolo Francione,⁹ Rocio Gallego-Durán,¹⁰ Javier Ampuero,¹⁰ Grazia Pennisi,¹¹ Rocio Aller,¹² Dina Tiniakos,^{5,13} Alastair Burt,⁵ Ezio David,¹⁴ Fabio Vecchio,¹⁵ Marco Maggioni,¹⁶ Daniela Cabibi,¹⁷ Duncan McLeod,¹⁸ Maria Jesus Pareja,¹⁹ Marco Y W Zaki,^{5,20,21} Antonio Grieco,²² Per Stål,^{23,24} Stergios Kechagias,²⁵ Anna Ludovica Fracanzani,^{9,26} Luca Valenti ^{26,27}, Luca Miele ^{6,7,8}, Piero Fariselli ³, Mohammed Eslam,²⁸ Salvatore Petta ¹¹, Hannes Hagström ^{23,24}, Jacob George ²⁸, Jörn M Schattenberg,^{2,29} Manuel Romero-Gómez ¹⁰, Quentin Mark Anstee,^{5,30} Elisabetta Bugianesi ¹

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2023-330815>).

For numbered affiliations see end of article.

Correspondence to

Dr Angelo Armandi, Division of Gastroenterology and Hepatology, Department of Medical Sciences, University of Turin, Turin, Italy; angelo.armandi@unito.it

AA and TS contributed equally.

AA and TS are joint first authors.

Received 1 August 2023

Accepted 22 December 2023

Published Online First

10 January 2024



Watch Video

gut.bmj.com/



► <http://dx.doi.org/10.1136/gutjnl-2023-331848>



© Author(s) (or their employer(s)) 2024. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Armandi A, Sanavia T, Younes R, et al. *Gut* 2024;**73**:825–834.

ABSTRACT

Objective Hyperferritinaemia is associated with liver fibrosis severity in patients with metabolic dysfunction-associated steatotic liver disease (MASLD), but the longitudinal implications have not been thoroughly investigated. We assessed the role of serum ferritin in predicting long-term outcomes or death.

Design We evaluated the relationship between baseline serum ferritin and longitudinal events in a multicentre cohort of 1342 patients. Four survival models considering ferritin with confounders or non-invasive scoring systems were applied with repeated five-fold cross-validation schema. Prediction performance was evaluated in terms of Harrell's C-index and its improvement by including ferritin as a covariate.

Results Median follow-up time was 96 months. Liver-related events occurred in 7.7%, hepatocellular carcinoma in 1.9%, cardiovascular events in 10.9%, extrahepatic cancers in 8.3% and all-cause mortality in 5.8%. Hyperferritinaemia was associated with a 50% increased risk of liver-related events and 27% of all-cause mortality. A stepwise increase in baseline ferritin thresholds was associated with a statistical increase in C-index, ranging between 0.02 (lasso-penalised Cox regression) and 0.03 (ridge-penalised Cox regression); the risk of developing liver-related events mainly increased from threshold 215.5 µg/L (median HR=1.71 and C-index=0.71) and the risk of overall mortality from threshold 272 µg/L (median HR=1.49 and C-index=0.70). The inclusion of serum ferritin thresholds (215.5 µg/L and 272 µg/L) in predictive models increased the performance of Fibrosis-4 and Non-Alcoholic Fatty Liver Disease Fibrosis Score in the longitudinal risk assessment of liver-related events (C-indices>0.71) and overall mortality (C-indices>0.65).

Conclusions This study supports the potential use of serum ferritin values for predicting the long-term prognosis of patients with MASLD.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Serum hyperferritinaemia is associated with liver disease severity and fibrosis in metabolic dysfunction-associated steatotic liver disease (MASLD). Preliminary evidence suggested that serum ferritin levels might have prognostic implications, but no robust evidence has been so far achieved.

WHAT THIS STUDY ADDS

⇒ Baseline hyperferritinaemia is associated with 50% risk of liver-related events and 27% of all-cause mortality, with a stepwise increase from values 215.5 µg/L and 272 µg/L, respectively. The inclusion of ferritin improves the performance of Fibrosis-4 Score (FIB-4) and Non-Alcoholic Fatty Liver Disease Fibrosis (NFS) Score for the prediction of liver-related events and all-cause mortality.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ These results suggest a longitudinal role for serum ferritin values in the risk stratification of MASLD patients.

INTRODUCTION

Serum ferritin assessment is part of the investigational panel to rule out other causes of liver damage in patients with suspected metabolic dysfunction-associated steatotic liver disease (MASLD),¹ formerly known as non-alcoholic fatty liver disease (NAFLD). Hyperferritinaemia, in the absence of significant iron overload, is detected in up to 30% of patients with MASLD.² As an acute-phase protein, serum ferritin is also actively synthesised during inflammatory processes or in the setting of

increased intracellular oxidative stress. Hyperferritinaemia has been widely reported in low-grade metabolic systemic inflammation conditions such as visceral obesity and type 2 diabetes mellitus (T2DM).³ Large cross-sectional studies have demonstrated a significant association of increased ferritin with the severity of liver fibrosis in patients with MASLD,⁴⁻⁶ especially in men and in those with a genetic predisposition to iron accumulation.⁷ Although normal ferritin levels, or less elevated, may reasonably exclude the presence and severity of liver fibrosis, serum ferritin on its own seems to lack overall accuracy in distinguishing between patients with or without severe or advanced liver fibrosis.⁸

In addition, the role of ferritin in the prediction of long-term outcomes in patients with MASLD is still uncertain. Hyperferritinaemia has been associated with increased mortality in patients with decompensated liver disease and patients on the liver transplantation waiting list.^{9,10} In a single-centre longitudinal study with 16 years of follow-up, high ferritin levels in patients with MASLD have been related to increased mortality, even after adjusting for confounders at baseline (age, sex, smoking, body mass index (BMI), diabetes, hypertension, cardiovascular disease and fibrosis stage). However, from the 713 patients with MASLD initially included, ferritin levels were only available in 222 cases.¹¹

The main aim of this study was to test the potential predictive value of serum ferritin on hepatic as well as extrahepatic outcomes and all-cause death among individuals with biopsy-proven MASLD. In particular, we evaluated the role of both the standard definition of hyperferritinaemia and different ferritin thresholds in the prediction models of hepatic/extrahepatic outcomes and all-cause death.

MATERIALS AND METHODS

We performed a multicentre cohort study of biopsy-confirmed MASLD patients from tertiary centres in Italy (Turin, Milan, Rome, Palermo), UK (Newcastle upon Tyne), Spain (Seville), Germany (Mainz), Sweden (Linköping and Stockholm) and Australia (Sydney). Inclusion criteria were the diagnosis of

MASLD confirmed by liver biopsy, age ≥ 18 years and available baseline value of serum ferritin. Ferritin levels of 300 $\mu\text{g/L}$ for men and 200 $\mu\text{g/L}$ for women were considered the 'upper limit of normal' (ULN).^{7,12}

A flow chart for the selection of the study subjects is shown in figure 1. Liver biopsies were performed from 1974 to 2015 for the suspicion of significant liver injury or fibrosis, based on the presence of chronically-altered liver biochemistry or hepatic steatosis at ultrasound. All biopsies were locally read and judged appropriate for analysis (average length of 25 mm and minimum 10 portal tracts) by expert liver pathologists, using the Clinical Research Network scoring system for MASLD histological grading and fibrosis staging.¹³ A threshold of 5% of hepatocytes showing steatosis was required for the diagnosis of MASLD. Metabolic dysfunction-associated steatohepatitis (MASH) was diagnosed when the combined presence of steatosis, lobular inflammation and ballooning was detected. Significant fibrosis was defined as fibrosis stage ≥ 2 . Advanced fibrosis was defined as fibrosis stage ≥ 3 . No signs of iron accumulation were detected in histology.

Exclusion criteria were represented by other causes of liver disease, including alcohol-related liver disease (cut-offs for exclusion of alcohol-induced liver disease were ≤ 210 g/week in men and ≤ 140 g/week in women; the history of alcohol consumption was estimated from standard clinical interviews during both the first visit and follow-up, based on medical records); iron overload caused by genetic predisposition (haemochromatosis, which was excluded by selecting patients with transferrin saturation $\leq 50\%$ in men and $\leq 45\%$ in women); Wilson disease and other cholestatic or autoimmune liver diseases; drug-induced liver injury caused by hepatotoxic medications (including methotrexate, systemic steroids, valproic acid, tamoxifen); viral hepatitis caused by hepatitis C and hepatitis B viruses.

At the time of the liver biopsy, clinical, anthropometric and laboratory data had been collected, including full blood count, routine liver biochemistry and metabolic profiling including glucose, lipids and iron. T2DM was diagnosed based on fasting serum glucose level ≥ 125 mg/dL or on active treatment. At the

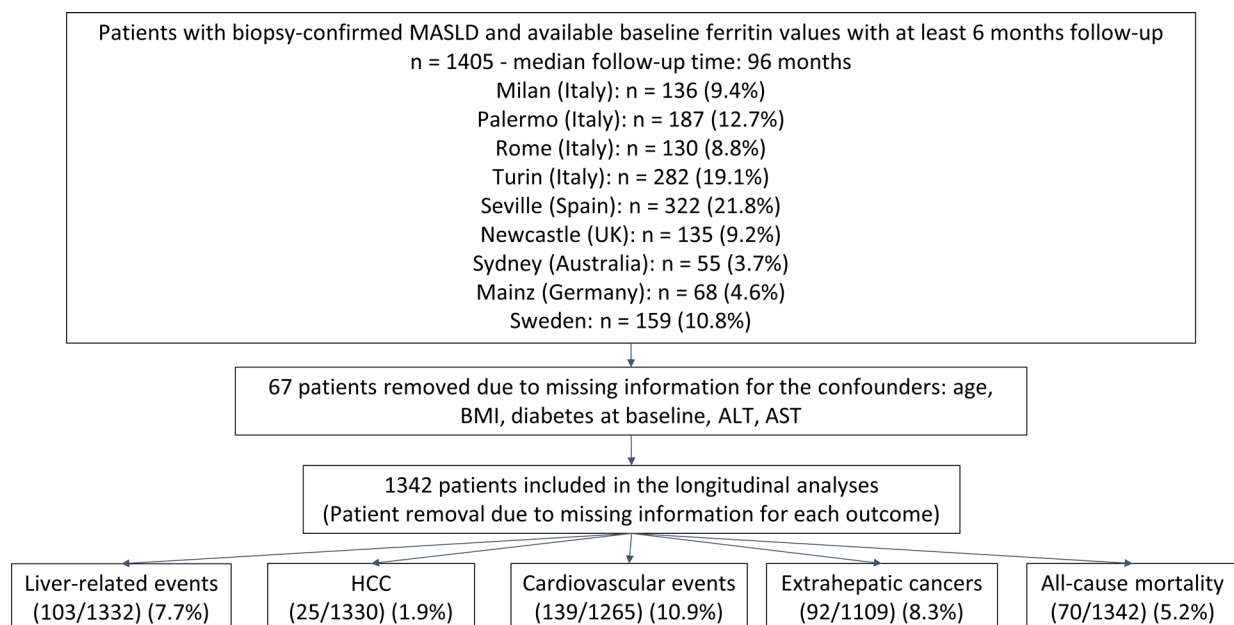


Figure 1 Flow chart of the study. ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; HCC, hepatocellular carcinoma; MASLD, metabolic dysfunction-associated steatotic liver disease.

Table 1 Demographic, clinical and histological characteristics of the patient population as a whole and subdivided into a 'high serum ferritin' (HSF) and a 'normal-low serum ferritin' (NLSF) group†

Group	Total (1474)	NLSF (1036)	HSF (438)	P value
Age (years)				
Median (SEMed)	48 (0.6)	46 (0.4)	51 (0.8)	<0.001*
Sex				
Male	952 (65%)	650 (63%)	302 (69%)	0.027*
Female	522 (35%)	386 (37%)	136 (31%)	
Body mass index (kg/m ²)				
Median (SEMed)	29.2 (0.2)	29.1 (0.2)	29.3 (0.3)	0.036*
Type 2 diabetes	393 (27%)	260 (25%)	133 (30%)	0.047*
ALT (IU/L)				
Median (SEMed)	62 (1.4)	59 (1.4)	74 (3.1)	<0.001*
AST (IU/L)				
Median (SEMed)	38 (0.6)	36 (0.4)	44 (1.9)	<0.001*
Haemoglobin (g/L)				
Median (SEMed)	148 (0.4)	147 (1)	150 (1)	0.0016*
Platelets (×10 ⁹ /L)				
Median (SEMed)	232 (2.1)	238 (2.5)	217 (3.1)	<0.001*
Albumin (g/L)				
Median (SEMed)	45 (0.01)	45 (0.01)	45 (0.2)	0.13
Total bilirubin (mg/dL)				
Median (SEMed)	196.2 (5.4)	185.4 (3.6)	216 (3.6)	<0.001*
Alkaline phosphatase (IU/L)				
Median (SEMed)	79 (1.4)	83 (1.6)	71 (1.6)	<0.001*
Gamma-glutamyltransferase (IU/L)				
Median (SEMed)	66 (2.1)	66 (2.3)	67 (4.3)	0.54
Glucose (mg/dL)				
Median (SEMed)	95.4 (0.7)	95.4 (0.7)	97.2 (1.8)	0.37
HDL (mg/dl)				
Median (SEMed)	21.6 (0.4)	23.4 (0.4)	21.6 (0.4)	0.028*
Triglycerides (mg/dL)				
Median (SEMed)	142.2 (3.6)	142.2 (3.6)	151.2 (5.4)	0.055
Total cholesterol (mg/dL)				
Median (SEMed)	91.8 (0.7)	91.8 (1.8)	93.6 (1.8)	0.092
MASH	902 (61%)	620 (60%)	282 (64%)	0.12
All stages fibrosis (F0 vs F1234)	1053 (71%)	716 (69%)	337 (77%)	0.0029*
Significant fibrosis (F01 vs F234)	607 (41%)	366 (35%)	241 (55%)	<0.001*
Advanced fibrosis (F012 vs F34)	297 (20%)	179 (17%)	118 (27%)	<0.001*

Data are expressed in terms of median and standard error of the median for continuous variables, or as number (%) for categorical variables. P values correspond to Kolmogorov-Smirnov and χ^2 tests for numerical and nominal data, respectively.

*Statistical significance, that is, p value <0.05.

†Serum ferritin levels of 300 µg/L for males and 200 µg/L for females, respectively.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; MASH, Metabolic-dysfunction associated steatohepatitis; SEMed, standard error of the median.

time of liver biopsy, none of the enrolled patients had clinical or radiological signs of decompensated cirrhosis. The baseline diagnosis of cirrhosis was exclusively based on histology.

For the longitudinal analysis, a minimum of 6 months of follow-up was recorded for each patient. For the Swedish cohort, longitudinal data were obtained from national registers. Medical visits were performed every 6 (for cirrhotic patients) or 12 months by hepatologists. Subjects whose health status was unknown for more than 12 months were considered lost at follow-up. Recorded events included liver-related outcomes (ascites, hepatic encephalopathy, oesophageal varices bleeding), the occurrence of hepatocellular carcinoma (HCC, defined by compatible radiological and/or histological findings), cardiovascular events (including acute coronary syndrome, peripheral arterial ischaemia, acute

cerebrovascular events), extrahepatic cancers (both epithelial and haematological neoplasia) and all-cause death.

MASLD treatment consisted of standard recommendations for lifestyle changes, including both increased physical activity and dietary changes.

Patients with histological evidence of cirrhosis underwent ultrasound surveillance for HCC and those who developed portal hypertension further underwent endoscopic screening for gastro-oesophageal varices, according to guidelines by liver societies.

Statistical analysis

Patients were divided into 'high serum ferritin' (HSF) and 'normal-low serum ferritin' (NLSF) groups, considering ferritin

Table 2 Median HRs of hyperferritinaemia (HSF vs NLSF groups) for long-term outcomes (multivariate Cox regression models)

Results	Liver-related events (103/1332)	HCC (25/1330)	Cardiovascular events (139/1265)	Extrahepatic cancer (92/1109)	Mortality (70/1342)
HSF versus NLSF HR (2.5%–97.5% CI)	1.531*† (0.955 to 2.458)	1.671* (0.634 to 4.44)	0.897* (0.576 to 1.396)	1.786*‡ (1.1 to 2.904)	1.267*† (0.701 to 2.29)

For each long-term outcome the table reports the median HRs and CI obtained on the training sets from the cross-validation of the multivariate Cox models for the serum ferritin values, categorised into HSF and NLSF groups. Confounders: age, sex, BMI, T2DM, AST and ALT. The ratio number of events/total number for each outcome is reported on top into brackets.

*Showing median log-rank test p value <0.05.

†Showing median Wald test p value of serum ferritin coefficient <0.1.

‡Showing median Wald test p value of serum ferritin coefficient <0.05.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; HCC, hepatocellular carcinoma; HSF, high serum ferritin; NLSF, normal-low serum ferritin; T2DM, type 2 diabetes mellitus.

levels of 300 µg/L for males and 200 µg/L for females. The baseline characteristics of these two groups were compared using χ^2 and Kolmogorov-Smirnov tests for nominal factors and numeric values, respectively. Longitudinal analysis was then performed to evaluate the ability of different ferritin values to discriminate the long-term incidence of the following outcomes, recorded at the end of the follow-up in each patient: liver-related events, HCC, cardiovascular events, extrahepatic cancer and overall mortality. Specifically, the evaluations were performed considering patients at: (1) HSF vs NLSF; (2) different cut-offs of ferritin values from 10th to 90th percentiles, defined from the entire patients' cohort. Cut-offs dividing the ferritin values into two classes with one reporting ≤ 10 events were excluded from evaluations.

The predictive power of the ferritin values concerning the long-term outcomes was assessed by applying a 5-time repeated 5-fold cross-validation of Cox proportional hazard regression and its regularised versions, that is, Lasso, Ridge and Elastic Net regressions.¹⁴ Regularisation is a method for preventing overfitting by controlling model complexity. It accomplishes this by penalising the coefficients of predictors that do not provide meaningful information to the model. Lasso penalisation allows for the generation of a sparse model (setting unimportant covariates to 0), while Ridge penalisation avoids the feature selection set by lasso by encouraging a grouping effect of the covariates used in the model. Finally, Cox regression combined with Elastic Net penalisation combines both lasso and ridge penalty terms to identify the more representative covariates in each group that contribute most to modelling the outcome. All the penalty terms were determined through cross-validation.

These four models were applied considering both ferritin and the following confounders: sex, age at biopsy, BMI, baseline T2DM, alanine aminotransferase (ALT) and aspartate aminotransferase (AST). For inclusion in the prediction model, each confounder with continuous values was binarised using the following thresholds: age at 50 years, BMI at 30 kg/m², ALT at 50 IU/L, AST at 36 IU/L following common clinical guidelines. Sex was excluded as a confounder in the evaluation of the groups of patients differentiated according to ULN, since the sex of the patients already conditions the state-of-the-art thresholds used to define these groups. To avoid the model overfitting, in each training set the penalty regularisation hyperparameters were set using 5-fold cross-validation. Splits of the dataset were prepared to have at least one event at either low or high ferritin values, keeping a balanced distribution of the events. Harrell's C-index was used to assess the model performance. Specifically, the same training/test folds were used to compare the C-index obtained by the full model with respect to the C-index from the same model using only the confounders as variables. All C-indices resulting from the cross-validation

were then compared using a one-sided Wilcoxon signed rank test. In addition, for each training/test fold, the difference between these C-indices was calculated and will be referred to as ΔC -index in the following. A positive ΔC -index indicates a gain in the prediction performance of the model when serum ferritin is considered a covariate.

Finally, we evaluated whether the ferritin can be informative according to the estimated Hazard Ratio (HR) and increase the prediction performance in terms of ΔC -index when combined with the non-invasive scoring systems Fibrosis-4 (FIB-4)¹⁵ and NAFLD Fibrosis Score (NFS),¹⁶ which are used to stratify patients with MASLD who are at risk of advanced fibrosis. Specifically, these two scores were calculated at the time of liver biopsy according to their originally reported formula:

► FIB-4: (age [years] × AST [U/L]) / ([platelets (× 10⁹/L)] × √ALT [U/L]).

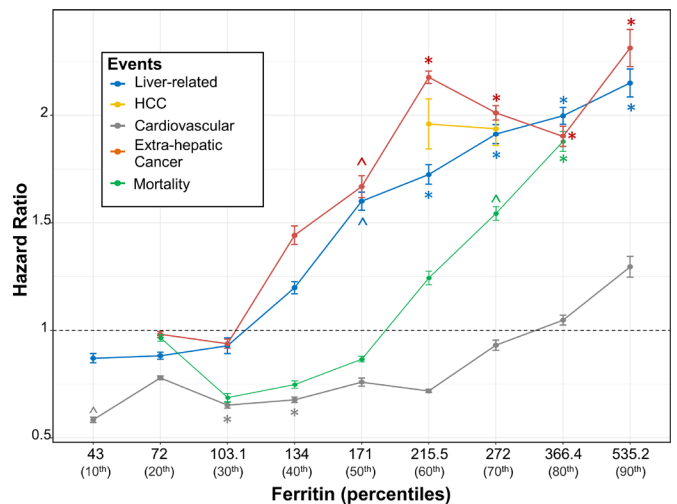


Figure 2 HRs of increasing serum ferritin thresholds for long term outcomes (multivariate Cox regression models). For each long-term outcome, the figure reports the average HRs and standard errors obtained on the training sets from the cross-validation of the multivariate Cox models, considering different thresholds for serum ferritin levels, ranging between the 10th and the 90th percentiles (x-axis). Confounders of the models: age, sex, BMI, T2DM, AST and ALT. All the obtained models showed a statistically significant log-rank test p value (<0.01). *Showing median Wald test p value of serum ferritin coefficient <0.05, ^Showing median Wald test p value of serum ferritin coefficient <0.1. ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; HCC, hepatocellular carcinoma; T2DM, type 2 diabetes mellitus.

Table 3 Multiple models testing prediction of long-term outcomes by hyperferritinaemia (HSF vs NLSF groups)

Method	Liver-related events (103/1332)	HCC (25/1330)	Cardiovascular events (139/1265)	Extrahepatic cancer (92/1109)	Mortality (70/1342)
Cox regression	0.68±0.007*	0.859±0.01*	0.716±0.014	0.667±0.024†	0.699±0.026†
Lasso-penalised Cox regression	0.664±0.008	0.818±0.017	0.689±0.007	0.668±0.031	0.694±0.024*
Ridge-penalised Cox regression	0.668±0.006*	0.842±0.01*	0.712±0.011	0.686±0.017*	0.651±0.026*
Elastic Net-penalised Cox regression	0.67±0.005*	0.832±0.019†	0.718±0.01	0.677±0.023	0.669±0.026*

Median Harrell's C-indices with corresponding standard errors of the median estimated on the test sets of the cross-validation. Cox proportional hazard model and its penalised variants (i.e., Lasso, Ridge and Elastic Net) were applied, considering serum ferritin values categorised into HSF and NLSF. Confounders: age, sex, BMI, T2DM, AST and ALT. The ratio number of events/total number for each outcome is reported on top into brackets.

*Reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value <0.05).

†P value <0.1 on the same test.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; HCC, hepatocellular carcinoma; HSF, high serum ferritin; NLSF, normal-low serum ferritin; T2DM, type 2 diabetes mellitus.

► NFS: $1.675 + 0.037 \times \text{age (years)} + 0.094 \times \text{BMI (kg/m}^2) + 1.13 \times \text{IFG/diabetes (yes=1, no=0)} + (0.99 \times \text{AST/ALT ratio}) (0.013 \times \text{platelet } [\times 10^9/\text{L}]) (0.66 \times \text{albumin [g/dl]})$.

The analyses were performed in the R environment (V.4.1.2), using the main R packages survival (V.3.2–13), glmnet (v. 4.1–4) and WRS2 (V. 1.1–4).

RESULTS

Baseline characteristics of the study cohort

A total of 1474 patients were considered. The baseline characteristics of the whole study population subdivided into a HSF and NLSF groups according to ULN thresholds are displayed in [table 1](#). Hyperferritinaemia was found in 438 patients (29.7%), who were mostly males, older, with a higher prevalence of obesity and T2DM compared with the NLSF group. MASH was diagnosed in 61% of cases with no difference between the two groups (64.0% at high ferritin vs 60.4%, p=0.12). Overall, advanced fibrosis was found in 20% of the patients. A higher proportion of patients with HSF had liver fibrosis (any stage) (77% vs 69%, p=0.0029), as well as significant fibrosis (≥F2) (55% vs 35%, p<0.001) and advanced fibrosis (F3-4) (27% vs 17%, p<0.001) ([table 1](#)).

Long-term outcomes of the study cohort

The minimum follow-up time for the whole cohort was 6 months (range 6–540, median 96 months) ([figure 1](#)). Sixty-seven patients were excluded for missing information on confounders (age, BMI, T2DM, ALT and AST). Hence, longitudinal analyses were performed on 91% of the subjects (1342/1474). During the follow-up, 7.7% (103/1332) of them developed liver-related events, 1.9% (25/1330) HCC, 10.9% (139/1265) had cardiovascular complications and 8.3% (92/1109) were diagnosed with extrahepatic cancers. Overall, 5.2% (70/1342) of the patients died.

Longitudinal analysis: HRs of HSF for long-term outcomes

According to current guidelines, serum ferritin values $\geq 300 \mu\text{g/L}$ for males and $\geq 200 \mu\text{g/L}$ for females were the threshold levels set for the definition of HSF versus NLSF groups. Cox model and its penalised versions were trained and tested with 5-fold cross-validation, repeating the procedure five times. For the preliminary evaluation of the main associations between HSF and long-term outcomes, we considered the median HRs and CIs for the serum ferritin values observed from the multivariable Cox regression models applied to the training datasets of the cross-validation. [Table 2](#) reports the results obtained from

the models in the training sets. Except for cardiovascular events, median HRs higher than 1.26 for all long-term outcomes and all-cause death were observed, with significant Wald test for extrahepatic cancer, and p values <0.1 for liver-related events and mortality. It is worth pointing out that hazards for HCC showed a high variability due to the lower number of events available (25 in total).

Longitudinal analysis: HRs of increasing serum ferritin thresholds for long-term outcomes

We then investigated the hazards' changes according to increasing ferritin thresholds by percentiles. [Figure 2](#) shows the average HRs from the Cox models considering the training splits of the cross-validation at ferritin values ranging from 10th to 90th percentiles. Online supplemental figure 1 and online supplemental table 1 report the median hazards for each long-term outcome. Due to the low number of events, the performance of serum ferritin on HCC prediction was evaluated only at two ferritin thresholds (215.5 and 272 $\mu\text{g/L}$), that included at least more than 10 events at either high or low ferritin values.

Except for cardiovascular events, HRs for long-term outcomes showed a stepwise increase by increasing ferritin thresholds, which was confirmed by statistical significance through the median p values from the Wald test applied to the corresponding ferritin coefficient ([figure 2](#)). Specifically, ferritin values above the 60th percentile (215.5 $\mu\text{g/L}$) were significantly associated with median HRs >1.70 for liver-related events and extrahepatic cancers, which achieved a median HR=2.11 at the 90th percentile (535.2 $\mu\text{g/L}$). For the overall mortality, serum ferritin values above 272 $\mu\text{g/L}$ (70th percentile) were associated with a median HR >1.40, reaching a median HR=1.82 at 366.4 $\mu\text{g/L}$ (80th percentile), showing statistical significance ([figure 2](#) and online supplemental tables 1 and 2). It was not possible to test the mortality at the 90th percentile due to the low number of events. However, the trend of the HR values at lower ferritin thresholds suggests that higher HRs are expected at ferritin values >366.4 $\mu\text{g/L}$.

Performance of serum ferritin in predicting long-term outcomes in patients with MASLD

To evaluate the contribution of serum ferritin (either as HSF or at different thresholds by percentiles) to multivariate prediction models, we compared the 'full' models including serum ferritin and the corresponding 'null' models using only confounders in terms of C-index calculated on the test sets from the cross-validation. The variations between these C-indices were then

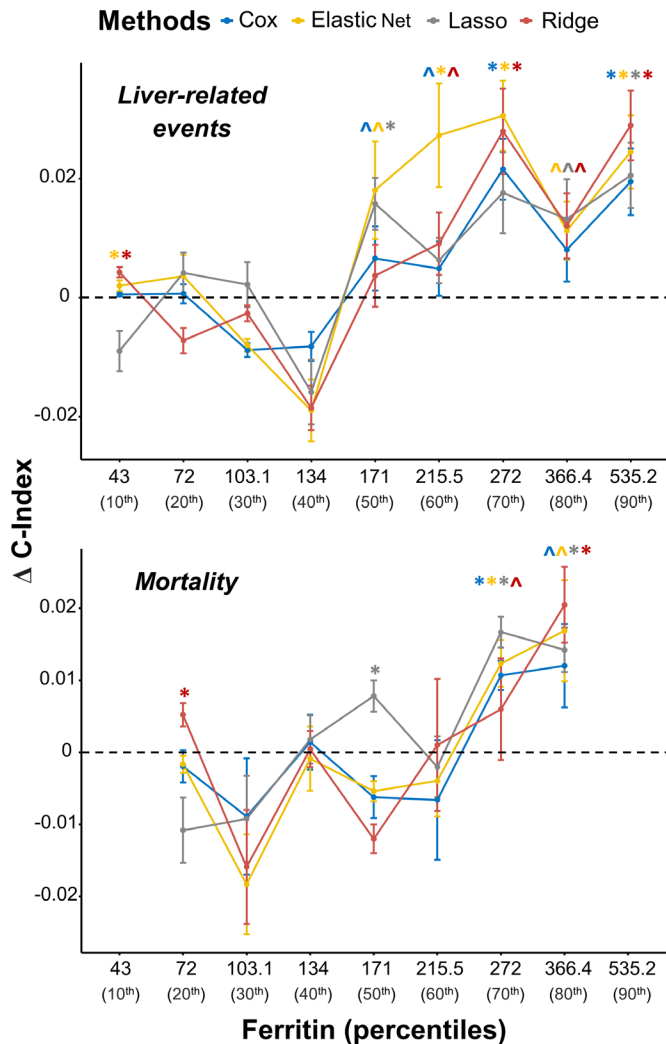


Figure 3 Variations of Harrell's C-indices measuring the predictive role of serum ferritin at different cut-offs with respect to liver-related events and mortality. The difference between the Harrell's C-indices obtained by the full model including the serum ferritin with respect to those obtained by the null model with only confounders is reported in the y-axis (Δ C-index). Cox proportional hazard model and its penalised variants (ie, Lasso, Ridge and Elastic Net) were applied, considering serum ferritin values at different cut-offs, corresponding to the range 10th - 90th percentiles (x-axis). Confounders: age, sex, BMI, T2DM, AST and ALT. *Reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value <0.05); ^p value <0.1 on the same test. Colours are displayed according to the methods reported in the legend. ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; T2DM, type 2 diabetes mellitus.

statistically compared using a one-sided Wilcoxon signed rank test.

When the ULN-defined groups were considered (table 3), three (out of four) prediction models for the liver-related events (ie, Cox without regularisation, Ridge and Elastic Net) achieved median C-indices between 0.67 and 0.68 showing a statistically significant increase of the predictive performance with respect to the null model (range 0.01–0.02). The highest C-indices were observed for HCC, with two prediction models (Cox and Ridge regression) achieving median C-indices of 0.86 and

0.84, respectively (average increase with respect to the corresponding null model of 0.01 and 0.02). Likewise, for mortality, all penalised models resulted statistically significant with median C-indices ranging between 0.65 and 0.69, respectively (average increase of 0.01). For the extrahepatic cancers, hyperferritinaemia showed a significant impact in the prediction of this outcome only in one case (when Ridge-penalised Cox model was applied), achieving a median C-index equal to 0.69 (average Δ C-index=0.01). Confirming the previous analysis, no incremental prediction performance of serum ferritin was observed for cardiovascular events.

Then, we repeated the same evaluation for serum ferritin at increasing percentiles cut-offs to investigate consistent trends of the C-index variations by increasing serum ferritin thresholds (figure 3). Specifically, for liver-related events, an incremental gain in terms of C-index was observed starting from ferritin values at the 50th percentile (171 μ g/L). From this threshold onwards, a statistical increase in C-index, ranging between 0.02 (with lasso-penalised Cox regression) and 0.03 (with ridge-penalised Cox regression), was achieved at serum ferritin levels of 272 μ g/L and particularly at 535.2 μ g/L, the highest ferritin threshold (90th percentile). An incremental gain in terms of prediction was also observed for overall mortality, becoming statistically significant for three models at the 70th percentile (272 μ g/L). At 80th percentile the Δ C-index resulted statistically significant for both Cox and its Ridge-penalised version but had higher variability (figure 3), due to the low number of events available. However, for the other models, the p values are <0.1.

Considering the other outcomes, reported in online supplemental figure 2, the most consistent gain in the prediction performance for extrahepatic cancers was observed above the 60th percentile of serum ferritin values (215.5 μ g/L) with statistically significant Δ C-indices from all four prediction models, but with variable trends of the Δ C-indices at higher thresholds compared to liver-related events and mortality. High variability was also observed for HCC, caused by the limited number of events, despite C-indices ranging between 0.79 and 0.88 (online supplemental table 2). However, the inclusion of serum ferritin as a covariate did not change the high prediction performance of the additional confounders considered in the model (i.e., sex, age, BMI, T2DM, ALT and AST). Finally, cardiovascular events are confirmed to be the least likely predicted by serum ferritin.

The inclusion of serum ferritin in models with non-invasive scoring systems can improve the prediction of liver-related events and death in patients with MASLD

We finally evaluated whether serum ferritin can increase the prediction of the long-term outcomes in combination with non-invasive scoring systems (FIB-4 and NFS) in terms of C-index. Observing the HRs in the training sets for the Cox model, serum ferritin was able to show a statistically significant impact in liver-related events, extrahepatic cancer and mortality with median HRs ≥ 1.2 for both scores (online supplemental tables 3 and 4), with the highest median HR observed for liver-related events (median HR=1.30 and =1.43 for FIB-4 and NFS, respectively). In terms of prediction performance on the test sets, we observed that, for both FIB-4 and NFS, the Δ C-index was statistically significant for predicting liver-related events in all four models and for predicting mortality in three out of the four models (tables 4 and 5).

Lastly, we also evaluated the risk distributions of liver-related events and mortality estimated by Cox model when we combined FIB-4 or NFS with serum ferritin values (figure 4). For each

Table 4 Multiple models testing prediction of long-term outcomes by serum ferritin combined with FIB-4 to predict long-term outcomes

Method	Liver-related events (103/1324)	HCC (23/1322)	Cardiovascular events (135/1254)	Extrahepatic cancer (94/1104)	Mortality (66/1334)
Cox regression	0.718±0.001*	0.825±0.008	0.57±0.007	0.544±0.012	0.652±0.018*
Lasso-penalised Cox regression	0.711±0.005*	0.829±0.008	0.627±0.006	0.512±0.011	0.632±0.013*
Ridge-penalised Cox regression	0.72±0.002*	0.834±0.009*	0.601±0.004	0.543±0.009	0.641±0.017*
Elastic net-penalised Cox regression	0.72±0.005*	0.829±0.008	0.639±0.006	0.514±0.011	0.63±0.016

Median Harrell's C-indices with corresponding standard errors of the median estimated on the test sets of the cross-validation. Cox proportional hazard model and its penalised variants (i.e., Lasso, Ridge and Elastic Net) were applied. The ratio number of events/total number for each outcome is reported on top into brackets.

*Reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value <0.05).

FIB-4, Fibrosis-4; HCC, hepatocellular carcinoma.

category of risk (low, medium and high), high ferritin values can significantly change the risk of either developing a liver-related event or dying, with predictive performance as reported in tables 4 and 5 (first row, corresponding to Cox model). The figure highlights thresholds commonly used for FIB-4 (1.3 and 2.67). It is worth noticing that, at low ferritin values (<100 µg/L), these thresholds correctly separate low, medium and high predicted risks. For higher ferritin values, the risk is influenced by the ferritin, and levels above 500 µg/L identified the great majority of subjects at high risk for liver-related events and overall mortality, even when FIB-4 is below 1.3 or NFS < -1.5. Figure 5 reports the different performance of FIB-4 risk categories in the prediction of liver-related events and mortality. Notably, the predictive performance of the high FIB-4 threshold for the risk of liver-related events and mortality is increased by the combination with the main predictive serum ferritin cut-offs (215.5 µg/L for liver-related events and 272 µg/L for mortality, respectively, highlighted in the previous analyses).

Therefore, serum ferritin can significantly improve the predictions of hard outcomes and mortality either alone or in combination with state-of-the-art non-invasive scoring systems.

DISCUSSION

In this large multicentre cohort study with a median follow-up of 96 months, we explored the potential role of baseline serum ferritin values in the prediction of long-term outcomes in patients with biopsy-proven MASLD. In keeping with previous studies, baseline hyperferritinaemia was found in almost one-third of our patients and was associated with higher stages of fibrosis. The longitudinal analysis demonstrates that (1) baseline hyperferritinaemia according to established thresholds is associated with a 50% increased risk of liver-related events and 27% of all-cause mortality; (2) a stepwise increase in baseline ferritin threshold is associated with a parallel increase in the risk of liver-related

events and overall mortality: the risk of developing liver-related events mainly increases from the threshold of 215.5 µg/L (on average, HR=1.71 and C-index=0.71) and the risk of overall mortality from the threshold 272 µg/L (on average, HR=1.49 and C-index=0.70); (3) the inclusion of these serum ferritin thresholds (215.5 µg/L and 272 µg/L) in predictive models can increase the performance of FIB-4 and NFS in the longitudinal risk assessment of liver-related events and overall mortality, respectively, and can help the identification of at-risk subjects even for FIB-4 or NFS values in the low-risk range.

Increased ferritin levels are commonly observed in a variety of clinical conditions, characterised by chronic inflammation and oxidative stress that impact iron metabolism, facilitating intracellular iron retention. Cancer and alcohol abuse are two frequent causes of hyperferritinaemia, but ferritin is also an acute phase reactant, and it is often increased in chronic inflammatory conditions such as in individuals with insulin resistance, T2DM and visceral obesity.¹⁷ This form of 'metabolic' hyperferritinaemia is caused by deregulated pathways in peripheral organs, in particular the adipose tissue.^{7, 18} Proinflammatory cytokines, such as tumour necrosis factor (TNF)-α or interleukin (IL)-1β upregulate the synthesis of ferritin. Hyperferritinaemia is found in up to 30% of patients with MASLD,¹⁹ where ferritin levels have been associated with the amount of hepatic fat accumulation and the severity of liver damage.⁵ Oxidative stress, derived from the accumulation of harmful intermediates of fatty acids from incomplete intracellular oxidation (lipotoxicity), may directly enhance ferritin transcription. As also observed in our cohort, this form of hyperferritinaemia may occur in the absence of increased intrahepatic iron deposition, being rather an epiphenomenon of metabolic dysfunction characterised by normal transferrin saturation and lack of iron overload at histology. Conversely, no signs of iron deficiency potentially affecting serum ferritin levels were observed in our cohort, as no patient had Hb levels below

Table 5 Multiple models testing prediction of long-term outcomes by serum ferritin combined with NFS to predict long-term outcomes

Method	Liver-related events (93/1179)	HCC (22/1177)	Cardiovascular events (123/1117)	Extrahepatic cancer (84/981)	Mortality (60/1189)
Cox regression	0.723±0.009*	0.885±0.002	0.682±0.005	0.604±0.001	0.65±0.02*
Lasso-penalised Cox regression	0.722±0.009*	0.884±0.003*	0.689±0.004	0.604±0.001	0.63±0.012*
Ridge-penalised Cox regression	0.722±0.009*	0.885±0.002	0.682±0.007	0.599±0.002	0.636±0.015*
Elastic net-penalised Cox regression	0.722±0.008*	0.884±0.002	0.689±0.004	0.603±0.001	0.63±0.011*

Median Harrell's C-indices with corresponding standard errors of the median estimated on the test sets of the cross-validation. Cox proportional hazard models and its penalised variants (i.e., Lasso, Ridge and Elastic Net) were applied. The ratio number of events/total number for each outcome is reported on top into brackets.

*Reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value <0.05).

†P value <0.1 on the same test.

HCC, hepatocellular carcinoma; NFS, NAFLD Fibrosis Score.

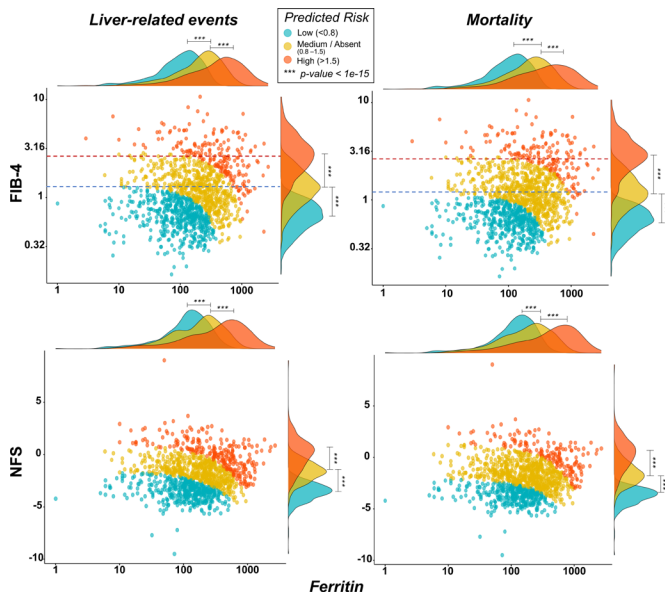


Figure 4 Distribution of predicted risk scores (low vs medium vs high) in liver-related events and mortality from Cox models considering the combination of FIB-4 and NFS with serum ferritin values as continuous variables. Colours are displayed according to the thresholds of the estimated relative risks reported in the legend. P values were obtained from Wilcoxon rank sum test. To improve the visualisation, log-scale was applied to Ferritin and FIB-4 values. Dashed lines highlight thresholds commonly used for FIB-4: 1.3 (blue line) and 2.67 (red line). FIB-4, Fibrosis-4; NFS, NAFLD Fibrosis Score.

100 g/L, the median age of the female population was 54 years (limiting the impact of periodic blood loss occurring during the fertile age) and no indirect signs of portal hypertension among the cirrhotic individuals were detected.

While the cross-sectional predictive value of serum ferritin has been explored in other large cohorts, our study is the first to carefully assess the potential predictive value of baseline ferritin levels on the long-term outcomes of MASLD patients.

Most studies on the prognostic role of hyperferritinaemia have been conducted on patients with decompensated cirrhosis,

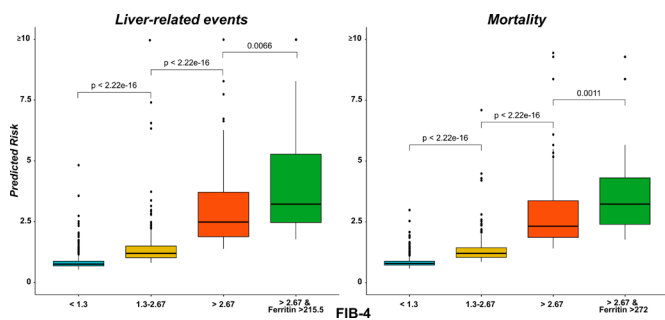


Figure 5 Boxplots highlighting the differences in the prediction risks for FIB-4 state-of-the-art threshold and in combination with serum ferritin cut-offs. State-of-the-art thresholds of FIB-4 (1.63 and 2.67) were initially analysed for the risk of liver-related events and mortality. The combinations of high FIB-4 threshold with the main predictive serum ferritin cut-offs (215.5 µg/L for liver-related events and 272 µg/L for mortality, respectively, highlighted in the previous analyses) display an increased predictive risk (green bar) compared with high FIB-4 threshold alone (orange bar). Wilcoxon rank sum test was applied to show the stepwise increase in prediction risk. FIB-4, Fibrosis-4.

where high ferritin values seem to be associated with increased mortality⁹ and with a worse short-term prognosis.¹⁰ In individuals with chronic liver disease, particularly due to alcohol and hepatitis C, hyperferritinaemia has been associated with liver-related mortality.²⁰ Only one study has been performed on biopsy-proven MASLD individuals so far, where after a follow-up of 16 years, hyperferritinaemia (>350 µg/L for males and 150 µg/L for females) was independently associated with increased mortality (HR 1.10 per year) regardless of different iron overload patterns.¹¹

Our study showed that the inclusion of ferritin levels in a multivariate model analysis significantly improves the predictive performance of clinical/biochemical variables such as age, sex, BMI, type 2 diabetes, AST and ALT. Of note, the risk of liver-related events doubled above the ferritin threshold of 535 µg/L (the 90th percentile in our cohort). Notably, this threshold is very close to the one (550 µg/L) emerged in previous studies in MASLD patients,^{21–23} in which patients seem to have benefit from iron depletion versus lifestyle changes alone in terms of improvement of liver enzymes.²³ Currently, 550 µg/L is the threshold proposed for grade 2 metabolic hyperferritinaemia, possibly associated with increased iron stores.⁷ Besides liver-related events and overall mortality, a consistent gain in the prediction of extracellular cancers was also observed above the ferritin threshold of 215 µg/L, consistent with the role of serum ferritin as a specific biomarker of cancers, well described in the literature; this aspect deserves attention as extrahepatic cancers represent the second cause of death in patients with MASLD.²⁴ Of note, serum ferritin was found to improve the predictive performance for liver-related events and mortality when combined with non-invasive scoring systems such as FIB-4 and NFS, thus showing that it is an important factor to be considered in the models related to these outcomes. Therefore, our study extends and supports the role of high ferritin in the predictions of hepatic complications and overall mortality, having an additional value on the clinical and biochemical variables usually included in algorithms for the non-invasive prediction of fibrosis and on the currently used non-invasive scores systems, namely FIB-4 and NFS.

The results of this study are supported by the elevated number of well-characterised biopsy-proven MASLD patients, prospectively enrolled and followed up in tertiary centres renowned for their expertise in MASLD, and by the robust analysis, validated by several models. Nevertheless, some limitations of this study need to be highlighted. The whole population represents a tertiary-level hospital referral, which might not reflect the same prevalence and the same characteristics when compared with the general population. The small amount of HCC developed during the follow-up may have affected the strength of the predictive models and limited the reliability of the results. In addition, data about the cause of death were missing, as well as the interval time between the event and the baseline biopsy.

In the landscape of MASLD, the availability of non-invasive prognostic biomarkers still represents an unmet need and liver histology remains the most reliable tool for prognostication.^{24,25} In this perspective, we extended the well-established association between hyperferritinaemia and fibrosis^{4–6} to a significant predictive performance of high ferritin levels for the risk of incident hard outcomes.

In conclusion, this longitudinal study strongly suggests that ferritin levels should be considered as an additional tool in the risk stratification of patients with MASLD, particularly in terms of liver-related events and overall mortality, and this finding deserves to be validated in prospective cohort studies.

Author affiliations

- ¹Division of Gastroenterology and Hepatology, Department of Medical Sciences, University of Turin, Turin, Italy
- ²Metabolic Liver Disease Research Program, I. Department of Internal Medicine, University Medical Center of the Johannes Gutenberg University, Mainz, Germany
- ³Computational Biomedicine Unit, Department of Medical Sciences, University of Turin, Turin, Italy
- ⁴Boehringer Ingelheim International GmbH, Ingelheim, Germany
- ⁵Translational & Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK
- ⁶Internal Medicine and Liver Transplant Unit, Department of Medical and Surgical Sciences, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
- ⁷CEMAD, Digestive Disease Center, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
- ⁸Department of Translational Medicine and Surgery, Università Cattolica del Sacro Cuore, Rome, Italy
- ⁹Unit of Medicine and Metabolic Disease, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
- ¹⁰UCM Digestive Diseases and SeLiver Group, Virgen del Rocio University Hospital, Institute of Biomedicine of Seville (HUVR/CSIC/US), University of Seville, Seville, Spain
- ¹¹Sezione di Gastroenterologia, PROMISE, Università di Palermo, Palermo, Italy
- ¹²Hospital Clínico de Valladolid, Valladolid, Spain
- ¹³Department of Pathology, Aretaieion Hospital, Medical School, National and Kapodistrian University of Athens, Athens, Greece
- ¹⁴Department of Pathology, Azienda Ospedaliero-Universitaria Città della Salute e della Scienza, University of Turin, Turin, Italy
- ¹⁵Dipartimento Universitario Medicina e Chirurgia Traslazionale, Università Cattolica del Sacro Cuore, Area Anatomia Patologica. Fondazione Policlinico Gemelli IRCCS, Rome, Italy
- ¹⁶Department of Pathology, Ca' Granda IRCCS Foundation, Milan, Italy
- ¹⁷Pathology Institute, PROMISE, University of Palermo, Palermo, Italy
- ¹⁸Department of Anatomical Pathology, Institute of Clinical Pathology and Medical Research (ICPMR), Westmead Hospital, Westmead, New South Wales, Australia
- ¹⁹Pathology Unit, Hospital Universitario de Valme, Seville, Spain
- ²⁰Biochemistry Department, Faculty of Pharmacy, Minia University, Minia, Egypt
- ²¹Centre for Research and Sustainability, Deraya University, New Minia, Minia, Egypt
- ²²Dipartimento Universitario Medicina e Chirurgia Traslazionale, Università Cattolica del Sacro Cuore, Area Medicina Interna, Gastroenterologia e Oncologia Medica, Fondazione Policlinico A. Gemelli IRCCS, Rome, Italy
- ²³Division of Hepatology, Department of Upper GI Diseases, Karolinska University Hospital, Stockholm, Sweden
- ²⁴Department of Medicine, Huddinge Karolinska Institutet, Stockholm, Sweden
- ²⁵Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden
- ²⁶Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy
- ²⁷Biological Resource Center Unit and Transfusion Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, Milan, Italy
- ²⁸Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital, University of Sydney, Westmead, New South Wales, Australia
- ²⁹Department of Medicine II, Saarland University Medical Center, Homburg, Germany
- ³⁰Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

Correction notice This article has been corrected since it published Online First. An affiliation has been added for author Marco Zaki.

X Angelo Armandi @AngeloArmandi, Dina Tiniakos @DinaTiniakos, Luca Valenti @lucavalenti75, Luca Miele @lmiele74, Hannes Hagström @hanneshagstrom and Manuel Romero-Gómez @mromerogomez

Contributors Manuscript concept and design: AA and EB. Writing: AA and TS. Data collection: RY, GPC, CR, OG, AL, PFrancione, RG-D, JA, GP, RA, DT, AB, ED, FV, MM, DC, DM, MJP, MYWZ, AG, PS, SK, ALF, LV, LM, ME, SP, HH, JG, JMS, MR-G and QMA. Statistical analysis: TS and PFariselli. Revision and editing: EB, QMA, MR-G, JMS, JG, HH, SP, LM, LV, ALF, SK, PS and PFariselli. Acceptance of the final version: all authors. EB is responsible for the overall content as guarantor.

Funding EB is supported by the Italian Ministry for Education, University and Research (MIUR) under the programme 'Dipartimenti di Eccellenza 2018-2022' Project code D15D18000410001. SP has received funding under PNRR M4C211.3 Heal Italia project PE00000019 CUP B73C22001250006, Italian PNRR-MAD-2022-12375656 project, Italian RF-2021-12372399 project and Italian PRIN-2022 2022L273C9 project. ALF is supported by: Ministry of Health grant Ricerca Finalizzata RF-2021-12374481 'Identification of lipid biomarkers to dissect hepatic and cardiovascular complication in NAFLD adult and children patients through a lipidomic approach'. JG is supported by the Robert W. Storr Bequest to the Sydney Medical Foundation, University of Sydney; a National Health and Medical Research Council of Australia (NHMRC) Program Grant (APP1053206), Project, Ideas and

Investigator grants (APP2001692, APP1107178, APP1108422, APP1196492) and a Cancer Institute, NSW grant (2021/ATRG2028).

Competing interests LM is supported by Investigator Driven Grants: Gilead, Intercept, Siemens Healthineers. Advisor/Consultancy: Alfa-Sigma, Boehringer-Ingelheim, BMS, Echosens, Galmed, Gilead Sciences, IBSA, Intercept, MEDA, MyGenomics, Merck Sharp & Dohme, Novartis, Pfizer, ProLon, Promethera, Resalis, Rottapharm-Madaus, Siemens Healthineers, Synageva. HH's institutions have received research funding from Astra Zeneca, EchoSens, Gilead, Intercept, MSD, Novo Nordisk and Pfizer. He has served as consultant or on advisory boards for Astra Zeneca, Bristol Myers-Squibb, MSD and Novo Nordisk and has been part of hepatic events adjudication committees for Boehringer Ingelheim, KOWA and GW Pharma.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study protocol was approved by appropriate regulatory bodies in all participating centers, and all patients gave their written informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs

Angelo Armandi <http://orcid.org/0000-0002-7245-4445>
 Tiziana Sanavia <http://orcid.org/0000-0003-3288-0631>
 Luca Valenti <http://orcid.org/0000-0001-8909-0345>
 Luca Miele <http://orcid.org/0000-0003-3464-0068>
 Piero Fariselli <http://orcid.org/0000-0003-1811-4762>
 Salvatore Petta <http://orcid.org/0000-0002-0822-9673>
 Hannes Hagström <http://orcid.org/0000-0002-8474-1759>
 Jacob George <http://orcid.org/0000-0002-8421-5476>
 Manuel Romero-Gómez <http://orcid.org/0000-0001-8494-8947>
 Elisabetta Bugianesi <http://orcid.org/0000-0002-0502-4381>

REFERENCES

- Rinella ME, Lazarus JV, Ratziu V, *et al.* A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol* 2023;79:1542–56.
- Valenti L, Dongiovanni P, Fracanzani AL, *et al.* Increased susceptibility to Nonalcoholic fatty liver disease in Heterozygotes for the Mutation responsible for hereditary Hemochromatosis. *Dig Liver Dis* 2003;35:172–8.
- Bozzini C, Girelli D, Olivieri O, *et al.* Prevalence of body iron excess in the metabolic syndrome. *Diabetes Care* 2005;28:2061–3.
- Kowdley KV, Belt P, Wilson LA, *et al.* Serum Ferritin is an independent predictor of histologic severity and advanced fibrosis in patients with nonalcoholic fatty liver disease. *Hepatology* 2012;55:77–85.
- Bugianesi E, Manzini P, D'Antico S, *et al.* Relative contribution of iron burden, HFE mutations, and insulin resistance to fibrosis in nonalcoholic fatty liver. *Hepatology* 2004;39:179–87.
- Buzzetti E, Petta S, Manuguerra R, *et al.* Evaluating the Association of serum Ferritin and hepatic iron with disease severity in non-alcoholic fatty liver disease. *Liver Int* 2019;39:1325–34.
- Valenti L, Corradini E, Adams LA, *et al.* Consensus statement on the definition and classification of metabolic hyperferritinaemia. *Nat Rev Endocrinol* 2023. 10.1038/s41574-023-00940-2 [Epub ahead of print 14 Dec 2023].
- Angulo P, George J, Day CP, *et al.* Serum ferritin levels lack diagnostic accuracy for liver fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 2014;12:1163–9.
- Meier JA, Bokemeyer A, Cordes F, *et al.* Serum levels of ferritin and transferrin serve as prognostic factors for mortality and survival in patients with end-stage liver disease: a propensity score-matched cohort study. *United European Gastroenterol J* 2020;8:332–9.
- Maiwall R, Kumar S, Chaudhary AK, *et al.* Serum Ferritin predicts early mortality in patients with decompensated cirrhosis. *J Hepatol* 2014;61:43–50.
- Hagström H, Nasr P, Bottai M, *et al.* Elevated serum ferritin is associated with increased mortality in non-alcoholic fatty liver disease after 16 years of follow-up. *Liver Int* 2016;36:1688–95.

- 12 Zoller H, Schaefer B, Vanclooster A. EASL clinical practice guidelines on haemochromatosis. *J Hepatol* 2022;77:479–502.
- 13 Kleiner DE, Brunt EM, Van Natta M, *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313–21.
- 14 Simon N, Friedman J, Hastie T, *et al.* Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw* 2011;39:1–13.
- 15 Sterling RK, Lissen E, Clumeck N, *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006;43:1317–25.
- 16 Angulo P, Hui JM, Marchesini G, *et al.* The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007;45:846–54.
- 17 Trombini P, Piperno A. Ferritin, metabolic syndrome and NAFLD: elective attractions and dangerous liaisons. *J Hepatol* 2007;46:549–52.
- 18 Moris W, Verhaegh P, Jonkers D, *et al.* Hyperferritinemia in nonalcoholic fatty liver disease: iron accumulation or inflammation? *Semin Liver Dis* 2019;39:476–82.
- 19 Dongiovanni P, Fracanzani AL, Fargion S, *et al.* Iron in fatty liver and in the metabolic syndrome: a promising therapeutic target. *J Hepatol* 2011;55:920–32.
- 20 Zhu Z, Yin J, Dawsey SM, *et al.* Association between serum Ferritin, incident primary liver cancer, and chronic liver disease mortality in the Linxian nutrition intervention trials: a nested case-control study. *J Gastroenterol Hepatol* 2021;36:3410–7.
- 21 Ryan JD, Armitage AE, Cobbold JF, *et al.* Hepatic iron is the major determinant of serum ferritin in NAFLD patients. *Liver Int* 2018;38:164–73.
- 22 Corradini E, Buzzetti E, Dongiovanni P, *et al.* Ceruloplasmin gene variants are associated with hyperferritinemia and increased liver iron in patients with NAFLD. *J Hepatol* 2021;75:506–13.
- 23 Valenti L, Fracanzani AL, Dongiovanni P, *et al.* Iron depletion by phlebotomy improves insulin resistance in patients with nonalcoholic fatty liver disease and hyperferritinemia: evidence from a case-control study. *Am J Gastroenterol* 2007;102:1251–8.
- 24 Sanyal AJ, Van Natta ML, Clark J, *et al.* Prospective study of outcomes in adults with nonalcoholic fatty liver disease. *N Engl J Med* 2021;385:1559–69.
- 25 Hagström H, Nasr P, Ekstedt M, *et al.* Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol* 2017;67:1265–73.

Supplementary Table 1. Median hazard ratios from multivariate Cox regression model associating serum ferritin levels to the long-term outcomes at different cut-offs. For each long-term outcome the table reports the median hazard ratios and confidence intervals (2.5% - 97.5%) obtained on the training sets from the cross-validation of the multivariate Cox models for the serum ferritin levels, considering different cut-offs corresponding to the range 10th - 90th percentiles. Confounders: age, sex, BMI, T2DM, AST and ALT. The ratio number of events/total number for each outcome is reported on top into brackets. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; T2DM, type 2 diabetes mellitus; HCC, hepatocellular carcinoma.

Ferritin Cut-off (percentiles)	Liver-related events (103/1332)	HCC (25/1330)	Cardiovascular events (139/1265)	Extrahepatic cancer (92/1109)	Mortality (70/1342)
43 (10%)	0.874 (0.424,1.812)	-	0.589 (0.342,1.009)	-	-
72 (20%)	0.852 (0.495,1.472)	-	0.785 (0.502,1.226)	0.955 (0.539,1.693)	0.941 (0.477,1.863)
103.1 (30%)	0.832 (0.517,1.341)	-	0.666 (0.444,0.999)	0.949 (0.565,1.597)	0.671 (0.385,1.169)
134 (40%)	1.219 (0.758,1.962)	-	0.646 (0.434,0.962)	1.378 (0.833,2.277)	0.779 (0.441,1.376)
171 (50%)	1.482 (0.934,2.358)	-	0.777 (0.514,1.171)	1.595 (0.985,2.583)	0.861 (0.486,1.524)
215.5 (60%)	1.708 (1.08,2.7)	1.803 (0.697,4.664)	0.721 (0.471,1.117)	2.15 (1.302,3.552)	1.22 (0.692,2.151)
272 (70%)	2.041 (1.272,3.262)	1.91 (0.737,5.018)	0.978 (0.618,1.547)	1.973 (1.196,3.256)	1.49 (0.83,2.674)
366.4 (80%)	2.04 (1.244,3.342)	-	1.009 (0.622,1.636)	1.855 (1.037,3.316)	1.821 (0.966,3.433)
535.2 (90%)	2.113 (1.182,3.78)	-	1.406 (0.774,2.553)	2.117 (1.163,3.854)	-

Supplementary Table 2. Testing prediction of ferritin on long-term outcomes at different cut-offs.

Median Harrell's C-indices with corresponding standard errors of the median were estimated on the test sets of the cross-validation. Cox proportional hazard model and its penalized variants (i.e. Lasso, Ridge and Elastic Net) were applied, considering serum ferritin values at different cut-offs, corresponding to the range 10th - 90th percentiles. Confounders: age, sex, BMI, T2DM, AST and ALT. The ratio number of events/total number for each outcome is reported on top into brackets. *: reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value ≤ 0.05); ^: p value < 0.1 on the same test. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; T2DM, type 2 diabetes mellitus; HCC, hepatocellular carcinoma.

Ferritin Cut-off (percentiles)	Method	Liver-related events (103/1332)	HCC (25/1330)	Cardiovascular events (139/1265)	Extrahepatic cancers (92/1109)	Mortality (70/1342)
43 (10%)	Cox	0.672±0.012	-	0.746±0.024*	-	-
	Lasso	0.664±0.016	-	0.744±0.028	-	-
	Ridge	0.666±0.013*	-	0.744±0.027*	-	-
	Elastic Net	0.662±0.016*	-	0.748±0.027*	-	-
72 (20%)	Cox	0.691±0.021	-	0.713±0.002^	0.605±0.014	0.740±0.025^
	Lasso	0.7±0.023	-	0.725±0.005	0.595±0.017	0.750±0.043
	Ridge	0.699±0.027	-	0.717±0.004	0.605±0.018	0.697±0.019
	Elastic Net	0.697±0.034	-	0.723±0.008	0.611±0.014	0.739±0.022
103.1 (30%)	Cox	0.695±0.016	-	0.714±0.024*	0.638±0.012	0.712±0.013
	Lasso	0.689±0.014	-	0.735±0.021^	0.623±0.005	0.698±0.013
	Ridge	0.693±0.022	-	0.731±0.027	0.632±0.012	0.710±0.012
	Elastic Net	0.689±0.014	-	0.727±0.023^	0.648±0.009	0.698±0.012
134 (40%)	Cox	0.669±0.011	-	0.672±0.022*	0.664±0.031	0.660±0.017
	Lasso	0.653±0.009	-	0.677±0.029	0.634±0.031	0.653±0.020
	Ridge	0.672±0.01	-	0.698±0.024	0.661±0.030	0.652±0.029
	Elastic Net	0.644±0.009	-	0.68±0.027	0.665±0.030	0.653±0.024
171 (50%)	Cox	0.677±0.003^	-	0.696±0.004	0.678±0.031*	0.722±0.010

	Lasso	0.689±0.007*	-	0.691±0.002	0.683±0.035	0.735±0.004*
	Ridge	0.684±0.005	-	0.696±0.006*	0.683±0.032*	0.711±0.013
	Elastic Net	0.689±0.004^	-	0.705±0.002*	0.664±0.033	0.702±0.017
215.5 (60%)	Cox	0.701±0.008^	0.785±0.004	0.72±0.002	0.686±0.042*	0.690±0.017
	Lasso	0.708±0.007	0.789±0.019	0.729±0.003	0.686±0.041*	0.659±0.020
	Ridge	0.705±0.007^	0.789±0.003	0.713±0.004	0.682±0.033*	0.698±0.019*
	Elastic Net	0.711±0.008*	0.787±0.02*	0.727±0.003*	0.686±0.039*	0.696±0.015^
272 (70%)	Cox	0.677±0.017*	0.877±0.04^	0.696±0.016	0.67±0.018*	0.692±0.027*
	Lasso	0.684±0.014	0.884±0.029	0.704±0.004	0.65±0.026	0.705±0.024*
	Ridge	0.686±0.015*	0.842±0.025	0.701±0.007	0.67±0.022*	0.708±0.024*
	Elastic Net	0.686±0.014*	0.832±0.035	0.707±0.008*	0.666±0.023*	0.703±0.030*
366.4 (80%)	Cox	0.714±0.014	-	0.694±0.009	0.628±0.006^	0.702±0.003*
	Lasso	0.72±0.016^	-	0.69±0.017	0.63±0.013	0.717±0.005^
	Ridge	0.723±0.018^	-	0.688±0.015	0.616±0.019*	0.713±0.009*
	Elastic Net	0.72±0.016^	-	0.69±0.015	0.621±0.012^	0.717±0.011^
535.2 (90%)	Cox	0.714±0.005*	-	0.677±0.012	0.657±0.009	-
	Lasso	0.713±0.003*	-	0.695±0.019	0.62±0.009	-
	Ridge	0.716±0.003*	-	0.68±0.017*	0.649±0.002	-
	Elastic Net	0.714±0.005*	-	0.688±0.015	0.642±0.007	-

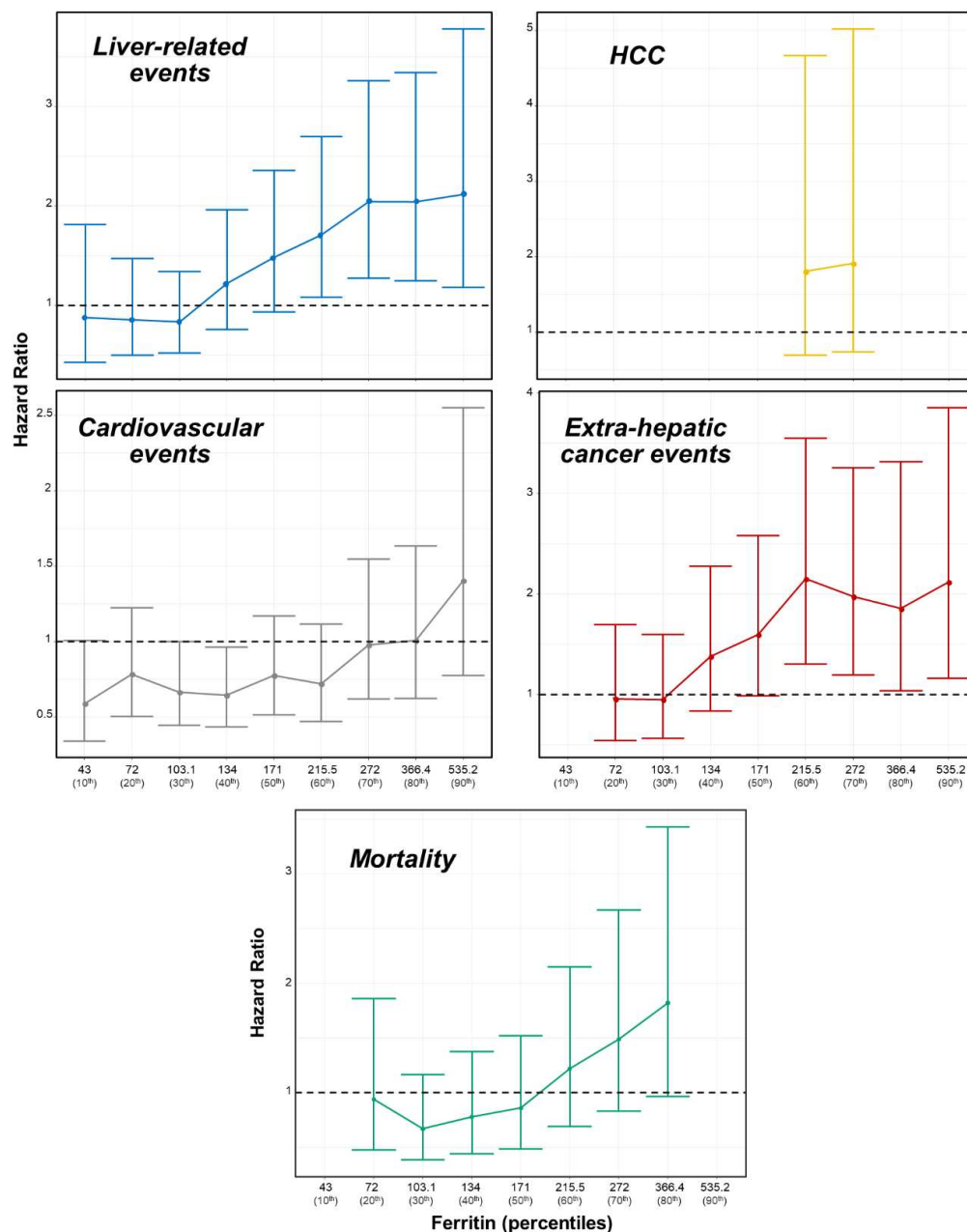
Supplementary Table 3. Median hazard ratios of serum ferritin combined with Fibrosis-4 (FIB-4) score to predict long term outcomes (multivariate Cox regression models). For each long-term outcome the table reports the median hazard ratios and confidence intervals (CI) obtained on the training sets from the cross-validation of the multivariate Cox models for the serum ferritin values. The ratio number of events/total number for each outcome is reported on top into brackets. †: showing median log-rank test p value <0.05 ; *: showing median Wald test p value for serum ferritin coefficient <0.05 . Abbreviations: HCC, hepatocellular carcinoma.

Results	Liver-related events (103/1324)	HCC (23/1322)	Cardiovascular events (135/1254)	Extrahepatic cancer (94/1104)	Mortality (66/1334)
Ferritin + FIB-4 Hazard Ratio (2.5%-97.5% CI)	1.297†* (1.104, 1.523)	1.007† (0.615, 1.646)	1.141† (0.951, 1.37)	1.291†* (1.093, 1.524)	1.17†* (0.934, 1.465)

Supplementary Table 4. Median hazard ratios of serum ferritin combined with NAFLD Fibrosis Score (NFS) to predict long term outcomes (multivariate Cox regression models). For each long-term outcome the table reports the median hazard ratios and confidence intervals (CI) obtained on the training sets from the cross-validation of the multivariate Cox models for the serum ferritin values. The ratio number of events/total number for each outcome is reported on top into brackets. †: showing median log-rank test p value <0.05 ; *: showing median Wald test p value for serum ferritin coefficient <0.05 , ^: showing median Wald test p value for serum ferritin coefficient <0.1 . Abbreviations: HCC, hepatocellular carcinoma.

Results	Liver-related events (93/1179)	HCC (22/1177)	Cardiovascular events (123/1117)	Extrahepatic cancer (84/981)	Mortality (60/1189)
Ferritin + NFS Hazard Ratio (2.5%-97.5% CI)	1.427†* (1.225, 1.663)	1.213† (0.805, 1.765)	1.123† (0.931, 1.354)	1.23†* (1.056, 1.434)	1.2†^ (0.945, 1.523)

Supplementary Figure 1. Median ferritin Hazard Ratios and confidence intervals from multivariate Cox regression models at different cut-offs. For each long-term outcome the figure reports the median hazard ratios and median confidence intervals obtained on the training sets from the cross-validation of the multivariate Cox models, considering different thresholds for the ferritin values, ranging between the 10th and the 90th percentiles (x-axis). Confounders of the models: age, sex, BMI, T2DM, AST and ALT. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; T2DM, type 2 diabetes mellitus, HCC, hepatocellular carcinoma.



Supplementary Figure 2. Variations of Harrell's C-index measuring the predictive role of ferritin values at different cut-offs with respect to HCC, extra-hepatic cancers and cardiovascular events. The difference between the Harrell's C-indices obtained by the full model including the serum ferritin with respect to those obtained by the null model with only confounders is reported in the y-axis. Cox proportional hazard model and its penalized variants (i.e. Lasso, Ridge and Elastic Net) were applied, considering serum ferritin values at different cut-offs, corresponding to the range 10th - 90th percentiles (x-axis). Confounders: age, sex, BMI, T2DM, AST and ALT. *: reporting statistically higher C-indices with respect to the same model tested on the same cross-validation split but considering only confounders, according to the one-sided Wilcoxon signed rank test (p value <0.05); ^: p value <0.1 on the same test. Colours are displayed according to the methods reported in the legend. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; T2DM, type 2 diabetes mellitus, HCC, hepatocellular carcinoma.

