Received: 25 August 2023 | Accepted: 27 November 2023

DOI: 10.1097/HEP.0000000000000707

ORIGINAL ARTICLE

OPEN

Combined use of the CLivD score and FIB-4 for prediction of liver-related outcomes in the population

Fredrik Åberg[1] | **Juho Asteljoki**[2,3,4] | **Ville Männistö**[5,6] | **Panu K. Luukkonen**[2,3,4]

[1]Transplantation and Liver Surgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

[2]Minerva Foundation Institute for Medical Research, Helsinki, Finland

[3]Department of Internal Medicine, University of Helsinki, Helsinki, Finland

[4]Abdominal Center, Helsinki University Hospital, Helsinki, Finland

[5]Department of Medicine, University of Eastern Finland, Kuopio, Finland

[6]Department of Medicine, Kuopio University Hospital, Kuopio, Finland

Correspondence
Fredrik Åberg, HUCH Meilahti Hospital, PB 372, HUS, Helsinki 00029, Finland.
Email: fredrik.aberg@helsinki.fi

Abstract

Background and Aims: A need exists for effective and practical tools to identify individuals at increased risk of liver-related outcomes (LROs) within the general population.

Approach and Results: We externally validated the chronic liver disease (CLivD) score for LROs in the UK Biobank cohort. We also investigated the sequential combined use of CLivD and fibrosis-4 (FIB-4) scores. Our analysis included 369,832 adults without baseline liver disease and with available data for CLivD and FIB-4 computation. LROs reflecting compensated or decompensated liver cirrhosis or HCC were ascertained through linkages with electronic health care registries. Discriminatory performance and cumulative incidence were evaluated with competing-risk methodologies. Over a 10-year follow-up, time-dependent AUC values for LRO prediction were 0.80 for $CLivD_{lab}$ (including gamma-glutamyltransferase), 0.72 for $CLivD_{non\text{-}lab}$ (excluding laboratory values), and 0.75 for FIB-4. $CLivD_{lab}$ demonstrated AUC values exceeding 0.85 for liver-related death and severe alcohol-associated liver outcomes. The predictive performance of FIB-4 increased with rising CLivD scores; 10-year FIB-4 AUC values ranged from 0.60 within the minimal-risk CLivD subgroup to 0.81 within the high-risk CLivD subgroup. Moreover, in the minimal-risk CLivD subgroup, the cumulative incidence of LRO varied from 0.05% to 0.3% across low-to-high FIB-4 strata. In contrast, within the high-risk CLivD subgroup, the corresponding incidence ranged from 1.7% to 21.1% (up to 33% in individuals with FIB-4 >3.25).

Conclusions: The CLivD score is a valid tool for LRO risk assessment and improves the predictive performance of FIB-4. The combined use of CLivD and FIB-4 identified a subgroup where 1 in 3 individuals developed LROs within 10 years.

Abbreviations: CLivD, predictive Liver Disease score; EASL, European Association for the Study of the Liver; FIB-4, fibrosis-4 index panel; GGT, gamma-glutamyltransferase; ICD-10, International Classification of Diseases Tenth Revision; LRO, liver-related outcome.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepjournal.com.

INTRODUCTION

Early detection of chronic liver disease is challenging but can be accomplished through noninvasive fibrosis testing in individuals who do not exhibit symptoms but are at a high risk for future liver-related outcomes (LROs). Currently, the recommended first-line method for risk stratification of fibrosis is the fibrosis-4 (FIB-4) index, primarily because of its widespread availability and low cost in clinical practice.[1,2] In other studies, a low FIB-4 value was used to exclude the risk of advanced liver fibrosis, whereas individuals with elevated FIB-4 values required further testing.[1,2]

However, the performance of the FIB-4 index is suboptimal when applied to unselected populations. For instance, in a large population study,[3] the sensitivity of an FIB-4 cutoff of 1.3 was only 37%. Therefore, the current EASL and AASLD guidelines emphasize that noninvasive tests such as FIB-4 should be used only in high-risk populations, specifically those with a sufficiently high pretest probability.[1,2] Nevertheless, a final definition of "high risk" in the general population has yet to be established.[4]

The Chronic Liver Disease (CLivD) score was recently introduced as a prognostic tool for predicting future LROs in the general population.[5] Calculated based on lifestyle factors alone or by including gamma-glutamyltransferase (GGT) as the only laboratory variable, the CLivD assigns a score on a continuous scale that corresponds to an individual's risk for LROs. The CLivD score can be used to identify high-risk populations, which necessitates subsequent evaluation with fibrosis testing; however, the effectiveness of such a strategy has not been firmly established. Although the CLivD score has been externally validated in 3 independent samples from Denmark, the United Kingdom, and the United States,[5,6] further validation studies are needed. In addition, studies on the combined or sequential use of the CLivD score and FIB-4 are lacking.

This study had 2 objectives. The first objective was to externally validate the predictive ability of the CLivD score for LROs and compare it with the predictive performance of FIB-4 in the UK Biobank population. The secondary objective was to assess the combined use of CLivD and FIB-4 scores for predicting LROs. We hypothesized that utilizing the CLivD score to define the at-risk population could improve the predictive performance of FIB-4. The findings of this study are important for designing pathways for screening liver fibrosis.

METHODS

This study used data from a UK Biobank sample. The UK Biobank is a community cohort study involving more than a half million individuals. Participants were interviewed between 2006 and 2010 at 22 assessment centers across England, Scotland, and Wales. All individuals aged 40–69 years and living within 25 miles of an assessment center (~9 million persons in total) were invited to participate. The participants completed a comprehensive health questionnaire, underwent a physical examination, and donated biological specimens. Follow-up data were obtained through record linkages with mortality, hospital admission, and cancer registries in the United Kingdom.[7] The UK Biobank Study was approved by the UK North West Multicentre Research Ethics Committee. Informed consent was obtained from all the participants. Some participants withdrew their consent.

Of all the participants (n = 502,412), we excluded those with known liver disease or liver cancer at or before baseline based on self-reported data and registry diagnoses, those with missing registry linkage, missing values in key baseline variables, baseline age <40 or >70 years, extreme outliers in the waist-hip ratio (<0.6 or >1.4), and those with chronic viral hepatitis (Figure 1).

Baseline variables

Alcohol consumption was evaluated by using a self-administered questionnaire. Participants were questioned about their drinking status and average weekly or monthly consumption of the number of glasses of red wine, champagne, white wine, pints of beer or cider, spirits, or glasses of fortified wine and other types of alcoholic drinks. For each participant, we calculated the mean consumption of ethanol in grams per day as described by Tavaglione et al,[8] assuming 2 units of pure alcohol in a pint of beer or cider, 1.5 units in a glass (125 mL) of red wine, champagne, white wine, fortified wine, and other alcoholic drinks, and 1 unit in a measure (25 mL) of spirit, where one unit contains 8 g of pure ethanol. Smoking status was categorized as never, previous, or current smoker, according to a self-

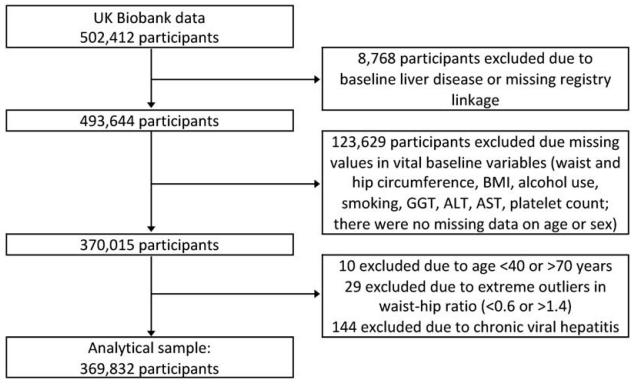

FIGURE 1 Flow chart of inclusion and exclusion criteria. Abbreviation: GGT, gamma-glutamyltransferase.

reported questionnaire. Body mass index and waist and hip circumference were measured at baseline. Diabetes at baseline was defined as either a nonfasting blood glucose ≥11.1 mmol/L, glycated hemoglobin ≥47 mmol/mol, a self-reported diabetes diagnosis or insulin treatment, or a registry code for diabetes. Laboratory methods used to generate biomarker data from blood samples have been described.[9]

Risk scores

The nonlaboratory $CLivD_{non\text{-}lab}$ score was calculated based on age, sex, smoking status (current vs. previous/never), weekly alcohol use, waist-hip ratio, and diabetes status (yes vs. no) as described.[5] The laboratory version of the $CLivD_{lab}$ also includes GGT (U/L). Both CLivD scores were categorized into subgroups (minimal, low, intermediate, and high risk) based on reported thresholds.[5] FIB-4 was calculated based on age, alanine aminotransferase, aspartate aminotransferase, and platelet count as described.[10] FIB-4 was categorized into low, intermediate, and high subgroups using a lower threshold of 1.3 if the age was <65, 2.0 if the age was ≥65, and a high threshold of 2.67. We also evaluated the lower FIB-4 threshold [1.3 for all ages (age-independent)].

Outcomes

The primary outcome was hospitalization, cancer, or death related to liver cirrhosis, or complications of cirrhosis (ie, "liver cirrhosis-related outcome event") defined by the International Classification of Diseases Tenth Revision (ICD-10) codes specified in Supplemental Table S1, http://links.lww.com/HEP/I150, in line with a recent consensus statement.[11] We also studied several secondary outcomes reflecting different types and severities of liver disease, as well as all-cause death (Supplemental Table S1, http://links.lww.com/HEP/I150). The follow-up time was calculated from the date of UK Biobank enrollment until the date of the outcome event, death, or censoring (end of the observation period or loss to follow-up), whichever occurred first. Follow-up was censored on the date of registry completion (December 1, 2021).

Statistical methods

To compare the groups, we used the chi-squared or Mann-Whitney tests, as appropriate. Predictive performance was evaluated using the competing-risk methodology. Time-dependent AUCs for the prediction of LROs were calculated from competing-risk regression models with the CLivD score or FIB-4 as the sole predictor, where death without liver disease was considered a competing-risk event. The cumulative incidence of LROs was calculated using the cumulative incidence function (Aalen-Johansen method). Time-dependent sensitivities, specificities, and positive and negative predictive values were estimated for various cutoff points, accounting for competing risks. Calibration was assessed by first dividing into deciles the predicted 10-year probabilities derived from cause-specific Cox regression. The mean predictive probability from cause-specific Cox regression was then plotted against the mean actual probability based on the competing-risk cumulative incidence function for each subgroup. Harrell's Adequacy Index was used to assess the fraction of new predictive information provided by a combination of the CLivD and FIB-4 scores compared with FIB-4 alone.[12] Statistical significance was defined as a two-tailed p value of <0.05. Data were analyzed using R software, version 4.0.0.

RESULTS

The final study cohort comprised 369,832 participants. Of the total, 48% were men, the mean age was 57 years, the mean body mass index was 27.2 kg/m^2, 5% had diabetes, and the mean alcohol use was 16.7 g of ethanol per day (Table 1). According to the $CLivD_{lab}$ scores, 25.6% belonged to the minimum-risk group, 69.6% to the low-risk group, 3.1% to the intermediate-risk group, and 1.6% to the high-risk group. According to the FIB-4 index, 64.9%, 32.9%, and 2.2% of patients belonged to the low-risk, intermediate-risk, and high-risk groups, respectively. No missing values were found for the parameters listed in Table 1.

During a median follow-up of 12.6 years (IQR: 12.0–13.4, 4,510,872 person-years), we found 1266 liver cirrhosis-associated outcome events (primary outcomes), 7671 other LROs (secondary outcomes), and 22,132 deaths without any liver-associated event (competing-risk events) (Table 2).

External validation of the CLivD score and comparison with FIB-4

The association between the CLivD score and rates of the primary outcome of cirrhosis-associated events was linear above a CLivD score of 0 (Supplemental Figure S1, http://links.lww.com/HEP/I150). Time-dependent AUC values at 10 years of follow-up for the prediction of the primary outcome were $CLivD_{lab}$, 0.80; $CLivD_{non\text{-}lab}$, 0.72; and FIB-4, 0.75 (Table 2). At 5 years, the corresponding AUC values were 0.81, 0.71, and 0.77, respectively. The AUC for the CLivD score remained stable throughout the follow-up period (Supplemental Figure S2, http://links.lww.com/HEP/I150).

TABLE 1 Baseline demographics

	All	Men	Women	*p*
Individuals	369,832	177,766	192,066	
Age	57 (8)	57 (8)	56 (8)	< 0.001
Body mass index (kg/m^2)	27.2 (4.6)	27.7 (4.1)	26.8 (5.0)	< 0.001
Waist circumference (cm)	90.1 (13.3)	96.7 (11.1)	84.0 (12.2)	< 0.001
Hip circumference (cm)	103.0 (8.8)	103.3 (7.4)	102.8 (10.0)	< 0.001
Waist-hip ratio	0.87 (0.09)	0.93 (0.06)	0.82 (0.07)	< 0.001
Diabetes	19,963 (5.4)	13,072 (7.4)	6891 (3.6)	< 0.001
Alcohol drinking status				< 0.001
Lifetime abstainer	19,185 (5.2)	5488 (3.1)	13,697 (7.1)	
Previous drinker	15,348 (4.1)	6816 (3.8)	8532 (4.4)	
Current drinker	335,299 (90.7)	165,462 (93.1)	169,837 (88.4)	
Alcohol intake (ethanol g/d)	16.7 (16.6)	22.3 (19.5)	11.4 (11.1)	< 0.001
Smoking habits				< 0.001
Never	200,043 (54.1)	86,639 (48.7)	113,404 (59.0)	
Previous	132,816 (35.9)	70,190 (39.5)	62,626 (32.6)	
Current	36,973 (10.0)	20,937 (11.8)	16,036 (8.3)	
Gamma-glutamyltransferase (U/L)	37 (40)	45 (46)	30 (31)	< 0.001
Alanine aminotransferase (U/L)	23 (14)	27 (15)	20 (12)	< 0.001
Aspartate aminotransferase (U/L)	26 (10)	28 (10)	24 (0)	< 0.001
$CLivD_{lab}$	0.25 (1.00)	0.62 (1.00)	−0.10 (0.87)	< 0.001
$CLivD_{non\text{-}lab}$	0.24 (0.95)	0.52 (1.01)	−0.03 (0.81)	< 0.001
FIB-4	1.35 (2.08)	1.43 (2.35)	1.28 (1.79)	< 0.001
$CLivD_{lab}$ risk group				< 0.001
Minimal	120,080 (32.5)	33,538 (18.9)	86,542 (45.1)	
Low	232,217 (62.8)	130,177 (73.2)	102,040 (53.1)	
Intermediate	11,487 (3.1)	9459 (5.3)	2028 (1.1)	
High	6048 (1.6)	4592 (2.6)	1456 (0.8)	
$CLivD_{non\text{-}lab}$ risk group				< 0.001
Minimal	94,849 (25.6)	32,258 (18.1)	62,591 (32.6)	
Low	257,262 (69.6)	129,892 (73.1)	127,370 (66.3)	
Intermediate	13,932 (3.8)	12,017 (6.8)	1915 (1.0)	
High	3789 (1.0)	3599 (2.0)	190 (0.1)	
FIB-4 risk group				< 0.001
Low	239,978 (64.9)	106,722 (60.0)	133,256 (69.4)	
Intermediate	121,649 (32.9)	65,502 (36.8)	56,147 (29.2)	
High	8205 (2.2)	5542 (3.1)	2663 (1.4)	

Note: Results are as mean (SD) or n (%).
Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

The discriminatory performance of $CLivD_{lab}$ 10-year AUC values for all secondary LROs was superior to that of FIB-4 (Table 2). The AUC values of $CLivD_{lab}$ were excellent for more severe forms of LROs, such as liver-associated death (AUC: 0.85–0.86), HCC (AUC: 0.82), alcohol-associated cirrhosis outcomes (AUC: 0.90), and alcohol-associated hepatitis (AUC: 0.85). The AUC values of $CLivD_{non\text{-}lab}$ were generally lower than those of FIB-4; however, $CLivD_{non\text{-}lab}$ outperformed FIB-4 in predicting alcohol-associated liver outcomes (Table 2).

The cumulative incidence of primary LRO at 10 years was 0.07% in the minimal-risk $CLivD_{lab}$ group, 0.2% in the low-risk group, 1.2% in the intermediate-risk group, and 4.8% in the high-risk group (Supplemental Figure S3, http://links.lww.com/HEP/I150). The corresponding 10-year cumulative incidence rates in the $CLivD_{non\text{-}lab}$ risk groups were 0.1%, 0.2%, 1.0%, and 2.7%, respectively (Supplemental Figure S3, http://links.lww.com/HEP/I150). Furthermore, the 10-year cumulative incidence rates in the low-risk, intermediate-risk, and high-risk FIB-4 groups were 0.1%, 0.3%, and 3.2%, respectively (Supplemental Figure

TABLE 2 AUC values at 10 years of follow-up of $CLivD_{lab}$, $CLivD_{non\text{-}lab}$, and FIB-4 for the prediction of the primary and secondary outcomes

		$CLivD_{lab}$	$CLivD_{non\text{-}lab}$	FIB-4
Primary outcome	Events	10-year AUC (95% CI)	10-year AUC (95% CI)	10-year AUC (95% CI)
Liver cirrhosis and cirrhosis-associated complications	1266	0.799 (0.799-0.799)	0.715 (0.715–0.716)	0.751 (0.750–0.751)
Secondary outcomes				
All liver outcomes	8937	0.651 (0.651–0.651)	0.600 (0.600–0.600)	0.589 (0.589–0.589)
Advanced liver disease	1859	0.760 (0.760–0.760)	0.689 (0.689–0.689)	0.707 (0.707–0.707)
Compensated liver cirrhosis	879	0.801 (0.801–0.801)	0.710 (0.709–0.710)	0.769 (0.769–0.770)
Decompensated liver cirrhosis	659	0.790 (0.790–0.790)	0.718 (0.718–0.719)	0.745 (0.745–0.745)
Primary liver cancer	248	0.745 (0.744–0.745)	0.686 (0.686–0.687)	0.672 (0.671–0.672)
HCC	142	0.822 (0.821–0.822)	0.733 (0.733–0.734)	0.747 (0.747–0.748)
Alcohol-associated cirrhosis	171	0.902 (0.902–0.903)	0.849 (0.848–0.849)	0.786 (0.786–0.787)
Alcohol-associated hepatitis	77	0.851 (0.850–0.852)	0.786 (0.786–0.787)	0.729 (0.729–0.730)
Unspecified alcohol-associated liver disease	260	0.847 (0.846–0.847)	0.780 (0.780–0.781)	0.751 (0.750–0.751)
Autoimmune liver disease	919	0.650 (0.650–0.651)	0.558 (0.558–0.559)	0.606 (0.605–0.606)
Primary biliary cholangitis	98	0.736 (0.735–0.737)	0.578 (0.578–0.579)	0.553 (0.552–0.554)
Cholangitis	736	0.646 (0.646–0.647)	0.592 (0.592–0.592)	0.611 (0.610–0.611)
Autoimmune hepatitis	93	0.585 (0.584–0.586)	0.539 (0.538–0.540)	0.625 (0.624–0.626)
Liver failure	376	0.661 (0.661–0.661)	0.634 (0.634–0.635)	0.598 (0.598–0.599)
Chronic hepatitis, not elsewhere classified	198	0.619 (0.618–0.619)	0.571 (0.570–0.571)	0.560 (0.560–0.561)
Fatty liver disease	3532	0.646 (0.646–0.646)	0.586 (0.586–0.586)	0.514 (0.514–0.514)
Alcoholic fatty liver	114	0.822 (0.821–0.823)	0.780 (0.780–0.781)	0.597 (0.597–0.598)
Unspecified liver disease	875	0.619 (0.618–0.619)	0.580 (0.579–0.580)	0.619 (0.619–0.619)
Liver disease as the primary cause of death	317	0.857 (0.856–0.857)	0.782 (0.782–0.782)	0.816 (0.816–0.817)
Liver disease as the primary or secondary cause of death	502	0.852 (0.852–0.853)	0.762 (0.761–0.762)	0.793 (0.793–0.793)
All-cause death	22,750	0.654 (0.654–0.654)	0.638 (0.638–0.638)	0.615 (0.615–0.615)

Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

S4, http://links.lww.com/HEP/I150). The 10-year cumulative incidence of the primary LRO in the 99.5th percentile of the $CLivD_{lab}$ score was 9.0%, and in the 99.5th percentile of the $CLivD_{non\text{-}lab}$ score was 3.3%.

The measures of discriminatory performance for various cutoffs of the $CLivD_{lab}$, $CLivD_{non\text{-}lab}$, and FIB-4 scores are shown in Table 3 and Supplemental Table S2, http://links.lww.com/HEP/I150. Calibration plots showed a slight underestimation of the observed 10-year risk at the higher end of the CLivD scores (Supplemental Figure S5, http://links.lww.com/HEP/I150).

Combined use of CLivD and FIB-4

When stratified by $CLivD_{lab}$ and FIB-4 risk groups, we found that within the minimal and low $CLivD_{lab}$ groups, the 10-year cumulative incidence of LROs remained low ($< 1.5\%$) in all FIB-4 risk groups (Figure 2 and Table 4). In contrast, within the high-risk $CLivD_{lab}$ group, the 10-year cumulative incidence varied from 1.7% with low FIB-4 values to 21.1% with high FIB-4 values (Figure 2 and Table 4). In contrast, $CLivD_{lab}$ could discriminate the risk of LROs within the low-risk, intermediate-risk, and high-risk FIB-4 groups (Supplemental Figure S6, http://links.lww.com/HEP/I150). These findings were similar for the combination of $CLivD_{non\text{-}lab}$ and FIB-4 (Supplemental Figure S7, http://links.lww.com/HEP/I150, Supplemenatl Figure S8, http://links.lww.com/HEP/I150).

We further stratified the study population into smaller subgroups based on 7 previously reported FIB-4 cutoff values, and 8 $CLivD_{lab}$ or 7 $CLivD_{non\text{-}lab}$ cutoff values, ensuring a minimum of 100 individuals per stratum. The number of individuals in each strata is shown in Supplemental Tables S3 and S4, http://links.lww.com/HEP/I150. We calculated the 10-year risk of LROs separately in each strata using the competing-risk cumulative incidence function. This enabled the identification of individuals with a 10-year risk of LROs of up to 33% (95% CI: 27%–39%) (Figure 3, CIs shown in Supplemental Tables S3 and S4, http://links.lww.com/HEP/I150).

The CLivD score modifies the predictive performance of FIB-4

The 10-year FIB-4 AUC values for predicting LROs increased from 0.60 among individuals in the minimal-

TABLE 3 Performance measures for $CLivD_{lab}$, $CLivD_{non\text{-}lab}$, and FIB-4 in predicting the primary liver-related outcome

Model and cutoff	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
$CLivD_{lab}$: minimal vs. low/intermediate/high (−0.258)	90.7	33.4	0.3	99.9
$CLivD_{lab}$: minimal/low vs. intermediate/high (2.066)	45.0	95.9	2.6	99.9
$CLivD_{lab}$: minimal/low/intermediate vs. high (2.784)	30.1	98.7	5.4	99.8
$CLivD_{non\text{-}lab}$: minimal vs. low/intermediate/high (−0.412)	88.7	26.3	0.3	99.9
$CLivD_{non\text{-}lab}$: minimal/low vs. intermediate/high (1.912)	26.4	95.7	1.5	99.8
$CLivD_{non\text{-}lab}$: minimal/low/intermediate vs. high (2.632)	10.6	99.2	3.0	99.8
FIB-4: low vs. intermediate/high (age-dependent; 1.3/2.0)	28.1	98.0	3.4	99.8
FIB-4: low vs. intermediate/high (1.3)	76.3	55.7	0.4	99.9

Note: Performance measures using additional cutoffs are shown in Supplemental Table S2, http://links.lww.com/HEP/I150
Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

risk $CLivD_{lab}$ group to 0.81 among those in the high-risk $CLivD_{lab}$ group (Table 4). Based on Harrell's Adequacy Index, a cause-specific Cox regression model with $CLivD_{lab}$ and FIB-4 provided 99.5% new prognostic information compared with a Cox model including FIB-4 as the only predictor (Supplemental Methods, http://links.lww.com/HEP/I150). Correspondingly, $CLivD_{non\text{-}lab}$ provided 98.8% of the new prognostic information compared to FIB-4 alone. In cause-specific Cox regression modeling, the interaction term between $CLivD_{lab}$ or $CLivD_{non\text{-}lab}$ and FIB-4 was significant ($p = 0.01$ and <0.001, respectively).

An improvement in the discriminatory performance of FIB-4 in individuals with higher CLivD scores was accompanied by a substantially increased sensitivity of the lower FIB-4 cutoff (Table 5). The specificity of this lower FIB-4 cutoff remained above 90% across CLivD subgroups when using the age-dependent FIB-4 cutoff (1.3/2.0), but decreased from 68% to 38% when using the FIB-4 cutoff of 1.3 for all ages (Table 5).

Exploratory analyses: a combination of $CLivD_{non\text{-}lab}$, FIB-4, and GGT

Given that GGT seemed to explain the improved performance of $CLivD_{lab}$ over $CLivD_{non\text{-}lab}$, we next explored a sequential approach, beginning with $CLivD_{non\text{-}lab}$, which

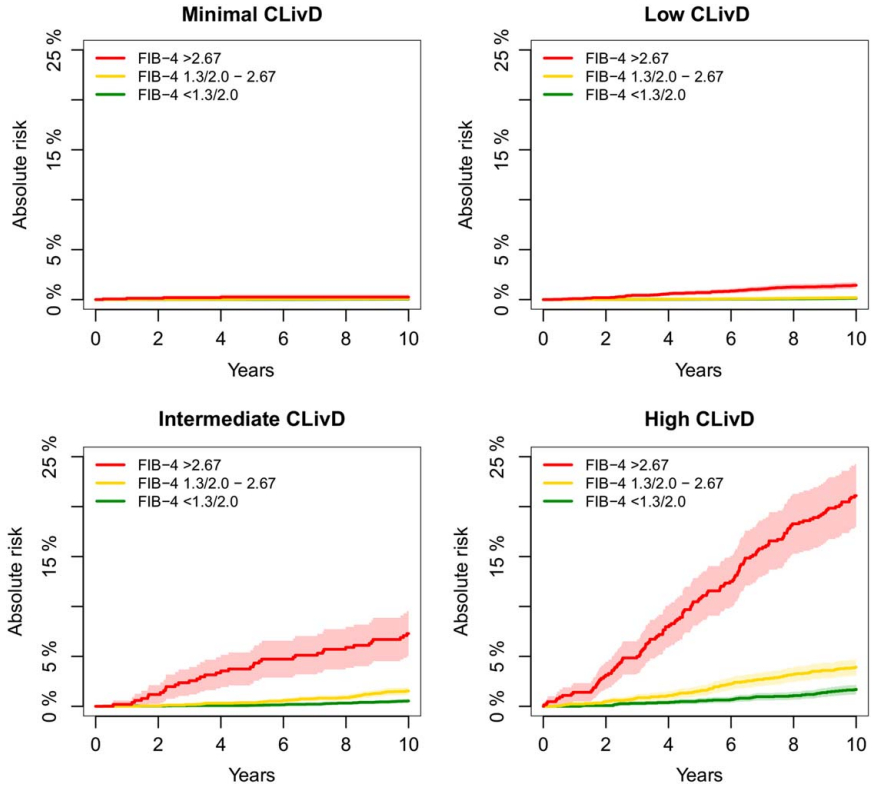

FIGURE 2 Cumulative incidence of the primary liver-related outcome according to FIB-4 within $CLivD_{lab}$ strata by the competing-risk cumulative incidence function (Aalen-Johansen method). Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

TABLE 4 AUC values at 10 years of follow-up of FIB-4 in predicting the primary liver-related outcome within CLivD risk groups

	Subjects	Outcome events	10-year AUC of FIB-4	10-year risk of liver-related outcome events		
				Low FIB-4 (%)	Intermediate FIB-4 (%)	High FIB-4 (%)
$CLivD_{lab}$						
Minimal risk	120,080	123	0.598 (0.593–0.603)	0.05	0.10	0.26
Low risk	232,217	627	0.663 (0.663–0.663)	0.13	0.19	1.43
Intermediate risk	11,487	179	0.741 (0.692–0.791)	0.54	1.54	7.28
High risk	6048	337	0.811 (0.783–0.838)	1.71	3.91	21.09
$CLivD_{non\text{-}lab}$						
Minimal risk	94,849	144	0.685 (0.684–0.686)	0.06	0.16	1.12
Low risk	257,262	791	0.728 (0.728–0.729)	0.13	0.24	2.64
Intermediate risk	13,932	203	0.767 (0.720–0.814)	0.47	1.08	8.31
High risk	3789	128	0.745 (0.688–0.802)	1.47	2.19	13.21

Note: Cumulative risks at 10 years by the Aalen-Johansen method are shown in subgroups based on the CLivD and FIB-4 scores.
Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

does not rely on laboratory data, followed by FIB-4 and GGT as separate variables. Compared with a cause-specific Cox regression model with $CLivD_{non\text{-}lab}$ and FIB-4 as predictors, a model with $CLivD_{non\text{-}lab}$, FIB-4, and GGT showed better discrimination (AUC: 0.78, 95% CI: 0.78–0.78 vs. 0.72, 95% CI: 0.72–0.72). The 10-year cumulative incidence rates of LROs were likewise consistently markedly higher across $CLivD_{non\text{-}lab}$ and FIB-4 strata among individuals with GGT ≥ compared to <120 U/L (Supplemental Table S5, http://links.lww.com/HEP/I150).

DISCUSSION

In this external validation study, the CLivD score version with GGT ($CLivD_{lab}$) demonstrated fairly similar discriminatory performance for incident LROs as the Finnish derivation study.[5] The competing-risk time-dependent AUC for $CLivD_{lab}$ was 0.80 in the UK Biobank cohort and 0.84 in the Finnish derivation cohort.

The absolute risks of LROs in the UK Biobank cohort were lower than those in the derivation cohort, which was an expected result because the derivation dataset used a broader set of ICD codes to define LRO. Here, we chose to define primary LRO according to a recent consensus on defining cirrhosis-associated events in epidemiological studies.[11] In addition, the mortality and cancer incidence rates in the UK Biobank sample were generally lower than those in the general population.[13]

This study compares the CLivD score with the FIB-4 index for the prediction of LROs in the population. The $CLivD_{lab}$ score demonstrated superior discriminatory performance for various types of LROs, particularly for more severe forms of liver disease such

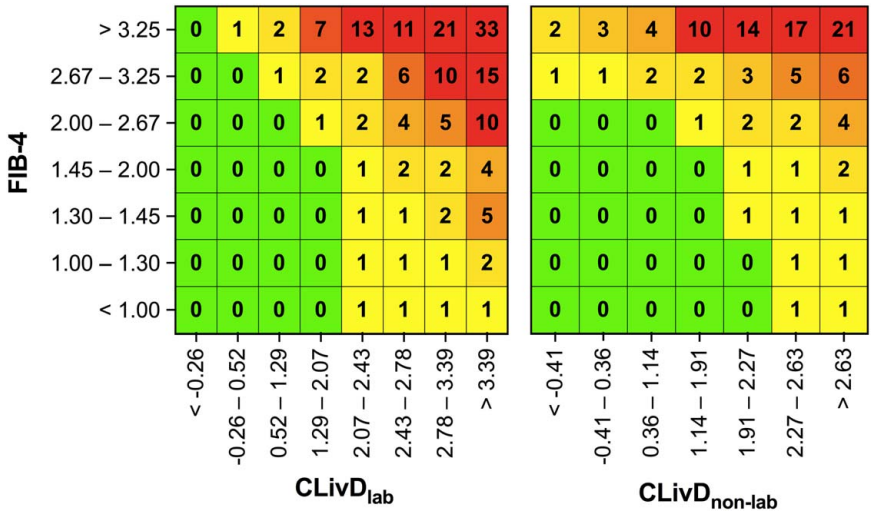

FIGURE 3 Cumulative incidence at 10 years of the primary liver-related outcome in subsets defined by several FIB-4 and CLivD score cutoffs. The cumulative incidences are calculated using the competing-risk cumulative incidence function (Aalen-Johansen method). Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

TABLE 5 Performance measures of FIB-4 for the primary liver-related outcome within specific CLivD risk groups

	FIB-4 (age-dependent, 1.3/2.0)				FIB-4 (1.3)			
	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
All individuals	28.1	98.0	3.4	99.8	76.3	55.7	0.4	99.9
Subgroup								
$CLivD_{lab}$ minimal	4.8	98.7	0.3	99.9	51.1	65.6	0.1	100
$CLivD_{lab}$ low	18.0	97.9	1.5	99.9	67.7	51.3	0.3	99.9
$CLivD_{lab}$ intermediate	27.7	95.9	8.0	99.0	84.0	44.7	1.9	99.5
$CLivD_{lab}$ high	50.5	91.4	24.1	97.2	93.1	37.7	7.5	99.0
$CLivD_{non\text{-}lab}$ minimal	16.0	98.6	1.2	99.9	64.9	65.6	0.2	99.9
$CLivD_{non\text{-}lab}$ low	26.6	98.0	2.8	99.8	74.5	52.9	0.4	99.9
$CLivD_{non\text{-}lab}$ intermediate	34.9	96.2	9.4	99.3	84.7	41.8	1.6	99.6
$CLivD_{non\text{-}lab}$ high	38.7	93.5	14.7	98.1	85.6	35.1	3.7	98.8

Abbreviations: CLivD, chronic liver disease; FIB-4, fibrosis-4.

as liver-associated death, HCC, and alcohol-associated cirrhosis. This is an important finding because alcohol is the leading cause of cirrhosis and LROs worldwide.[14] With only a fraction of harmful drinkers developing cirrhosis, there is a need for simple and inexpensive risk prediction tools suitable for use in this population.

The second aim of this study was to explore the sequential combination of the CLivD and FIB-4 scores to improve risk predictions. This combined use identified a subset of individuals with a 10-year LRO risk as high as 33% (Figure 3). These risk estimates were based on a cumulative incidence function with competing risks. The risk estimate for this high level is exceptional in general population studies. In individuals with such a high risk, proactive interventions to reduce this risk are most likely merited even without further extensive liver fibrosis testing.

The CLivD score modified the predictive performance of FIB-4 for LROs. Specifically, the performance of FIB-4 was poor in the subgroup with low CLivD scores, and good or excellent in those with high CLivD scores. This reflects the well-established spectrum effect[15] and is also in line with findings from the UK Biobank cohort, which reported an AUC value for FIB-4 of 0.59 in individuals without metabolic risk factors and 0.70–0.72 among those with metabolic risk factors[16]; however, that study did not assess alcohol use or competing risks. The effect modification of FIB-4 performance by the CLivD score reflects the need to accurately define the at-risk population that best benefits from liver disease risk stratification using FIB-4.

The strengths of our study include a large external validation sample size of nearly 400,000 individuals with a median 12.6-year follow-up, providing a robust dataset for analyses. We analyzed several types of LROs to confirm their consistency. The use of competing-risk methods and the cumulative incidence function instead of regression modeling are additional strengths.

Although large, the UK Biobank sample may not represent the general population or the diversity of populations worldwide,[13] potentially limiting the generalizability of our findings. Nonetheless, this limitation is more important for absolute risk estimates than it is for risk-factor associations.[17] In addition, the CLivD score has been externally validated in British, Danish, and US population cohorts.[5,6] We acknowledge that the LROs were based on registry diagnoses, which may not be fully accurate or complete. Nonetheless, the ICD codes used to define the primary outcome were selected based on a comprehensive systematic review, which further reported positive predictive values for cirrhosis ranging from 83% to 89% in external validation studies.[11]

Implications

Based on our findings, targeting the FIB-4 in individuals with higher CLivD scores enhances the discriminatory

performance of the FIB-4. Therefore, incorporating the CLivD score into liver disease risk stratification pathways could improve the identification of individuals at high risk of future LROs. Defining high-risk populations based on CLivD scores is an inexpensive means of targeting further evaluation and testing, thereby optimizing the allocation of resources and reducing unnecessary testing in low-risk individuals. The CLivD score is based on widely available variables and can be integrated into digital health care solutions, completed in general practice, or obtained through the Internet. Valid self-measurements of the waist-hip ratio can be acquired using simple measuring tapes or digital photography technology.[18,19]

$CLivD_{non\text{-}lab}$ does not require a single blood test, making it an appealing first-line test for large-scale screening. Although the overall performance of $CLivD_{non\text{-}lab}$ was poorer than that of $CLivD_{lab}$, the combined use of $CLivD_{non\text{-}lab}$ and FIB-4 could also identify a subgroup with a high 10-year risk for LROs. Adding GGT to FIB-4 in second-line testing improved risk prediction based on $CLivD_{non\text{-}lab}$.

GGT shows only a modest correlation with the severity of liver disease in cross-sectional studies.[20] However, GGT has been shown to be a stronger predictor of future liver disease compared to other liver enzymes.[5,21,22] GGT serves as a marker for hepatic and whole-body oxidative stress, potentially reflecting the mechanisms leading to disease development,[23] including the synergistic effects of alcohol and metabolic dysfunction.[24] Therefore, GGT can be viewed as an indicator of disease risk rather than existing liver disease.[23] Furthermore, GGT levels decrease with healthy lifestyle interventions,[25,26] emphasizing its potential as a monitoring tool, especially when incorporated into the CLivD score.

Further research is needed in diverse populations to validate the performance of the CLivD score and its combined use with FIB-4, and to explore its impact on patient outcomes and cost-effectiveness.

In conclusion, the CLivD score is a valuable tool for risk assessment of chronic liver disease, and the CLivD score improves the predictive performance of FIB-4. Incorporating CLivD scores into risk stratification pathways could enable the early detection of chronic liver disease.

AUTHOR CONTRIBUTIONS

Conceptualization: Fredrik Åberg, Juho Asteljoki, Ville Männistö, and Panu K. Luukkonen. Data curation: Juho Asteljoki and Fredrik Åberg. Formal analysis: Fredrik Åberg and Juho Asteljoki. Methodology: Fredrik Åberg, Juho Asteljoki, Ville Männistö, and Panu K. Luukkonen. Project administration: Panu K. Luukkonen. Software: Fredrik Åberg and Juho Asteljoki. Supervision: Panu K. Luukkonen and Ville Männistö. Visualization: Fredrik Åberg and Panu K. Luukkonen. Writing—original draft: Fredrik Åberg. Writing—review and editing: Juho Asteljoki, Ville Männistö, and Panu K. Luukkonen.

ACKNOWLEDGMENTS

This study was conducted using the UK Biobank Resource (application number 62797). Copyright © (2023), NHS England. Re-used with the permission of the NHS England and UK Biobank. All rights reserved. This work uses data provided by patients and collected by the NHS as part of their care and support. This research used data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (research which commenced between October 1, 2020 and March 31, 2021, grant ref MC_PC_20029; April 1, 2021 to September 30, 2022, grant ref MC_PC_20058). We would like to thank Editage (www.editage.com) for English language editing.

FUNDING INFORMATION

Fredrik Åberg was supported by the Academy of Finland (#338544), Sigrid Jusélius Foundation, and Wilhelm and Else Stockmann Foundation. Panu K. Luukkonen was supported by the Academy of Finland (350545) and the Sigrid Jusélius, Instrumentarium, Novo Nordisk, Orion, and Finnish Medical Foundations. The researchers are independent of the funders. The funding sources had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

CONFLICTS OF INTEREST

The authors have no conflicts to report.

ORCID

Fredrik Åberg https://orcid.org/0000–0002–3833–0705

REFERENCES

1. European Association for the Study of the Liver Clinical Practice Guideline Panel; Berzigotti A, Tsochatzis E, Boursier J, Castera L, Cazzagon N, Friedrich-Rust M, et al. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis—2021 update. J Hepatol. 2021; 75:659–89.
2. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology. 2023;77:1797–835.
3. Graupera I, Thiele M, Serra-Burriel M, Caballeria L, Roulot D, Wong GL-H, et al. Low accuracy of FIB-4 and NAFLD fibrosis scores for screening for liver fibrosis in the population. Clin Gastroenterol Hepatol. 2022;20:2567–76.
4. Åberg F, Jula A, Färkkilä M, Salomaa V, Erlund I, Männistö S, et al. Comparison of various strategies to define the optimal target population for liver fibrosis screening: A population-based cohort study. United European. Gastroenterol J. 2022;10:1020–8.

5. Åberg F, Luukkonen PK, But A, Salomaa V, Britton A, Petersen KM, et al. Development and validation of a model to predict incident chronic liver disease in the general population: The CLivD score. J Hepatol. 2022;77:302–11.
6. Song J, Jiang ZG. A good step toward low-cost prognostication of liver-related outcome awaits more validation. J Hepatol. 2022; 77:887–9.
7. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779.
8. Tavaglione F, De Vincentis A, Jamialahmadi O, Pujia R, Spagnuolo R, Picardi A, et al. Inborn and acquired risk factors for severe liver disease in Europeans with type 2 diabetes from the UK Biobank. JHEP Rep. 2021;3:100262.
9. UK Biobank. Biomarker assay quality procedures: Approaches used to minimise systematic and random errors (and the wider epidemiological implications). Version 1.2. 2019. Accessed August 2023. https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum_-biochemistry.pdf
10. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43:1317–25.
11. Shearer JE, Gonzalez JJ, Min T, Parker R, Jones R, Su GL, et al. Systematic review: development of a consensus code set to identify cirrhosis in electronic health records. Aliment Pharmacol Ther. 2022;55:645–57.
12. Harrell Jr FE. Regression Modeling Strategies. 2nd edn. Switzerland: Springer International Publishing; 2015. doi: 10.1007/978-3-319-19425-7.
13. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol. 2017;186: 1026–34.
14. Devarbhavi H, Asrani SK, Arab JP, Nartey YA, Pose E, Kamath PS. Global burden of liver disease: 2023 update. J Hepatol. 2023;79:516–37.
15. Usher-Smith JA, Sharp SJ, Griffin SJ. The spectrum effect in tests for risk prediction, screening, and diagnosis. BMJ. 2016; 353:i3139.
16. De Vincentis A, Tavaglione F, Jamialahmadi O, Picardi A, Antonelli Incalzi R, Valenti L, et al. A polygenic risk score to refine risk stratification and prediction for severe liver disease by clinical fibrosis scores. Clin Gastroenterol Hepatol. 2022;20: 658–73.
17. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: Prospective cohort study and individual participant meta-analysis. BMJ. 2020;368:m131.
18. Barrios P, Martin-Biggers J, Quick V, Byrd-Bredbenner C. Reliability and criterion validity of self-measured waist, hip, and neck circumferences. BMC Med Res Methodol. 2016;16:49.
19. Neufeld EV, Seltzer RA, Sazzad T, Dolezal BA. A multidomain approach to assessing the convergent and concurrent validity of a mobile application when compared to conventional methods of determining body composition. Sensors (Basel). 2020;20:E6165.
20. Petta S, Macaluso FS, Barcellona MR, Cammà C, Cabibi D, Di Marco V, et al. Serum γ-glutamyl transferase levels, insulin resistance and liver fibrosis in patients with chronic liver diseases. PLoS One. 2012;7:e51165.
21. Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Ability of noninvasive scoring systems to identify individuals in the population at risk for severe liver disease. Gastroenterology. 2020;158:200–14.
22. McLernon DJ, Donnan PT, Ryder S, Roderick P, Sullivan FM, Rosenberg W, et al. Health outcomes following liver function testing in primary care: A retrospective cohort study. Fam Pract. 2009;26:251–9.
23. Koenig G, Seneff S. Gamma-glutamyltransferase: A predictive biomarker of cellular antioxidant inadequacy and disease risk. Dis Markers. 2015;2015:818570.
24. Niemelä O. Biomarker-based approaches for assessing alcohol use disorders. Int J Environ Res Public Health. 2016;13:166.
25. St George A, Bauman A, Johnston A, Farrell G, Chey T, George J. Effect of a lifestyle intervention in patients with abnormal liver enzymes and metabolic risk factors. J Gastroenterol Hepatol. 2009;24:399–407.
26. Lazo M, Solga SF, Horska A, Bonekamp S, Diehl AM, Brancati FL, et al. Effect of a 12-month intensive lifestyle intervention on hepatic steatosis in adults with type 2 diabetes. Diabetes Care. 2010;33:2156–63.

How to cite this article: Åberg F, Asteljoki J, Männistö V, Luukkonen PK. Combined use of the CLivD score and FIB-4 for prediction of liver-related outcomes in the population. Hepatology. 2024;80:163–172. https://doi.org/10.1097/HEP.0000000000000707