JIM Original Article

doi: 10.1111/joim.13813

Performance of novel collagen turnover biomarkers to detect increased liver stiffness in MASLD

Hannes Hegmar[1,2], Thomas Wiggers[3], Patrik Nasr[4], Johan Vessby[5], Stergios Kechagias[4], Nils Nyhlin[6], Hanns-Ulrich Marschall[7,†], Åsa Danielsson Borssén[8], Rickard Strandberg[2], Morten Karsdal[3], Diana Julie Leeming[3], Mattias Ekstedt[4] & Hannes Hagström[1,2]

From the [1]Division of Hepatology, Department of Upper GI, Karolinska University Hospital, Stockholm, Sweden; [2]Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden; [3]Nordic Bioscience, Herlev, Denmark; [4]Department of Gastroenterology and Hepatology, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden; [5]Department of Medical Sciences, Gastroenterology Research Group, Uppsala University, Uppsala, Sweden; [6]Department of Gastroenterology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden; [7]Sahlgrenska University Hospital, Gothenburg, Sweden; and [8]Umeå University Hospital, Umeå, Sweden

Abstract. Hegmar H, Wiggers T, Nasr P, Vessby J, Kechagias S, Nyhlin N, et al. Performance of novel collagen turnover biomarkers to detect increased liver stiffness in MASLD. *J Intern Med.* 2024;**296**:177–86.

Background. Cleavage products from collagen formation and degradation hold potential as first-line biomarkers for the risk of advanced fibrosis in patients with metabolic dysfunction-associated steatotic liver disease (MASLD). Here, we evaluated the performance of PRO-C3, PRO-C6, C4M, PRO-C18L, and the clinical score ADAPT (age, diabetes, PRO-C3, and platelet count) to detect patients with an LSM >8 kPa or >12 kPa in comparison to the Fibrosis-4 Index (FIB-4).

Methods. Serum from patients with MASLD (n = 269) from six Swedish University Hospitals was analyzed using enzyme-linked immunosorbent assay-based methods. Liver stiffness measurement (LSM) by vibration-controlled transient elastography was performed. The area under the curve (AUC), calibration curves, and net benefit analysis were used.

Results. An LSM >8 kPa was found in 108 (40.1%) patients. PRO-C3, PRO-C6, C4M, and PRO-C18L had AUCs ranging from 0.48 to 0.62. ADAPT had the highest AUC (0.73, 95% confidence interval [CI] = 0.67–0.79) to detect patients >8 kPa, compared to FIB-4 (0.71, (95%CI = 0.64–0.77, p = 0.35), and had a higher net benefit compared to FIB-4 from a probability threshold of 15%. FIB-4 and ADAPT performed equally well to detect patients with an LSM >12 kPa, AUC 0.76 versus 0.76, p = 0.93.

Conclusions. ADAPT seems to be marginally better than FIB-4 in identifying patients with an LSM >8 kPa. However, the clinical utility of ADAPT as a first line test is uncertain, especially in low-risk populations. The overall performance of FIB-4 was similar to that of ADAPT in detecting patients with an LSM of >12 kPa. Altogether, the results suggest that ADAPT might be useful to detect earlier stages of fibrosis in MASLD, but that FIB-4 remains a first-line test for advanced fibrosis.

Keywords: NAFLD, MASLD, liver fibrosis, noninvasive tests

List of abbreviations: ADAPT, (age, diabetes, PRO-C3, platelet count); AGA, American Gastroenterological Association; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, Area under the curve; BMI, Body Mass Index; CAP, Controlled Attenuation Parameter; CI, Confidence Interval; FIB-4, fibrosis-4 index; ELISA, Enzyme-Linked Immunosorbent Assay; IQR, Interquartile Range; LSM, Liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; NAFLD, nonalcoholic fatty liver disease; VCTE, vibration-controlled transient elastography

Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) (previously nonalcoholic fatty liver disease) is the most common liver disease globally with an estimated prevalence of around 40% [1, 2], and encompasses a heterogeneous spectrum of disease stages.

Fibrosis stage is the most reliable predictor of liver-related outcomes [3]. Most clinical guidelines suggest a two-step approach to identify patients outside of hepatology services with a high risk for advanced liver fibrosis. The first step is a simple noninvasive test such as the Fibrosis-4 Index (FIB-4), [4]. Patients with values corresponding to intermediate or high risk of advanced fibrosis should undergo a second test, usually a liver stiffness measurement (LSM) with vibration-controlled transient elastography (VCTE). Patients with VCTE values above 8 kPa are considered at a higher risk of advanced fibrosis and are suggested to be referred to hepatology services [4]. However, the first step in this approach (FIB-4) lacks diagnostic accuracy [5].

Collagen cleavage products reflect fibrosis formation and degradation and may hold potential as biomarkers for fibrosis. PRO-C3, a neo-epitope of collagen III reflecting active fibrogenesis, can differentiate significant (fibrosis stage $\geq$2) and advanced (fibrosis stage $\geq$3) liver fibrosis from lower stages, both alone and in conjunction with clinical scores such as ADAPT (age, diabetes, PRO-C3, platelet count) [6, 7]. Here, we assessed the performance of collagen cleavage products compared to FIB-4 as biomarkers for the detection of elevated liver stiffness in patients with MASLD.

Material and methods

Study population

This cross-sectional study included patients with MASLD from an ongoing prospective cohort study from university hospitals in six regions in Sweden: Stockholm, Linköping, Uppsala, Umeå, Örebro, and Gothenburg. Steatosis was diagnosed using ultrasound, computed tomography, or magnetic resonance imaging. Most patients were referred from primary care for routine evaluation of elevated transaminases or incidental findings of hepatic steatosis. Patients with other etiologies of steatosis, including potentially steatogenic medications and a self-reported alcohol consumption exceeding 30 g per week in men and 20 g per week in women, were excluded.

Variables

The primary analysis considered if collagen-based biomarkers PRO-C3, PRO-C6, C4M, PRO-C18L, or ADAPT outperformed FIB-4 in detecting a liver stiffness >8 kPa. However, as most patients with a VCTE measurement of >8 kPa do not have advanced fibrosis, a secondary analysis was considered if the collagen-based biomarkers outperformed FIB-4 in detecting a liver stiffness of >12 kPa.

All participants underwent LSM by VCTE after fasting for a minimum of 3 h, using the M- or XL-probe as appropriate. Measurements with an LSM >7.1 kPa and an interquartile range (IQR) of the median >30% were considered unreliable [8]. Blood samples were collected at the same visit as VCTE or within 30 days of baseline.

Clinical parameters including height, weight, waist and hip circumference, medications, alcohol consumption, smoking status, and regular laboratory tests such as complete blood count, alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin, fasting glucose, hemoglobin A1c, and creatinine were collected. Type 2 diabetes was defined as a formal diagnosis, a fasting glucose >7.0 mmol/L, hemoglobin A1c >48 mmol/mol, or use of any antidiabetic medication. Hypertension was defined as blood pressure of or above 140/90 mmHg, a formal diagnosis, or use of an antihypertensive medication. Hyperlipidemia was defined as triglycerides >2.0 mmol/L, total cholesterol >5.2 mmol/L, low-density lipoprotein >3.0 mmol/L, or use of a lipid-lowering medication. Obesity was defined as a body mass index >30 kg/m^2. The ADAPT (age, diabetes, PRO-C3, platelet count) and FIB-4 (AST, ALT, age, platelet count) scores were calculated as previously described [7, 9].

Measurements of PRO-C3, PRO-6, C4M, and PRO-C18L

Each biomarker was assessed using an enzyme-linked immunosorbent assay at the Nordic Bioscience laboratory, Herlev, Denmark [10].

†Deceased.

The manuscript has been handled by an external editor: Senior Professor Olov Wiklund, Department of Molecular and Clinical Medicine at Institute of Medicine, Sahlgrenska University Hospital Gothenburg, Sweden

Statistical analysis

The data are presented as median values and IQR, or as total numbers and percentages. Statistical comparisons between groups were by the Mann–Whitney *U* test or Chi^2-test. The levels of each biomarker were compared between patients using two cut-offs: an LSM by VCTE ≤8 kPa and >8 kPa, and patients with an LSM ≤12 kPa and >12 kPa. Univariable logistic regression was performed for each cut-off and biomarker, and the area under the curve (AUC) was calculated. To illustrate the tests' performance in correctly identifying the majority of patients above the set cut-offs (8 or 12 kPa), we set the sensitivity to 90% and calculated test characteristics based on that level. We evaluated the currently suggested cut-offs 1.30 for FIB-4 and 4.45 for ADAPT [11], and as an explorative analysis, we evaluated ADAPT as a second-line test in patients with a FIB-4 >1.30, using a cut-off of 4.45.

We used logistic regression to calibrate the level of FIB-4 or ADAPT to correspond with an absolute risk (0%–100%) of an elevated liver stiffness of 8 or 12 kPa. We then assessed calibration of FIB-4 and ADAPT and performed a decision curve analysis to evaluate the net benefit of each model [12, 13]. To compare the number of FIB-4 and ADAPT, respectively, needed to avoid one unnecessary LSM by VCTE, the opt-out test trade-offs were calculated at different probability thresholds. It was calculated as $\frac{p}{1-p}/(NB_{test1} - NB_{test2})$, where p is the probability threshold and *NB* is the net benefit [13].

Ethical considerations

Ethical approval was obtained from the Swedish Ethical Review Authority (dnr 2016/2137-31), and patients provided written and oral consent.

RESULTS

We included 269 patients with MASLD (Fig. S1). Baseline characteristics are presented in Table 1. The median age was 55 years (43–64), and 50.3% were male.

Evaluation of test characteristics at the >8 kPa cut-off

The prevalence of an LSM >8 kPa was 40.1% (n = 108). PRO-C3 had the highest AUC among the examined individual biomarkers at 0.62 (95% confidence interval [CI] = 0.55–0.69). The other biomarkers showed comparatively lower AUCs: PRO-C6 0.51 (95%CI = 0.43–0.58), C4M 0.56 (95%CI = 0.49–0.63), and PRO-C18L 0.48 (95%CI = 0.41–0.55). The ADAPT score had an AUC of 0.73 (95%CI = 0.67–0.79), which was numerically but not statistically significantly higher than the FIB-4 AUC of 0.71 (95%CI = 0.64–0.77, p = 0.35). AUCs for the examined biomarkers are presented in Fig. 1a. When the cut-off of 1.3 for FIB-4 was used, the sensitivity and specificity was 63% and 67% to detect an LSM >8 kPa. The PPV and NPV were 56% and 73%, respectively. The suggested cut-off of 4.45 for the ADAPT score had sensitivity and specificity of 96% and 28%. The PPV and NPV was 47% and 92%, respectively (Table S1). The test characteristics of FIB-4 and ADAPT using these cut-offs to detect patients with an LSM >12 kPa are presented in Table S2.

When calculating the test characteristics at a sensitivity of 90%, the ADAPT score demonstrated the highest specificity (37%), followed by FIB-4 (30%).

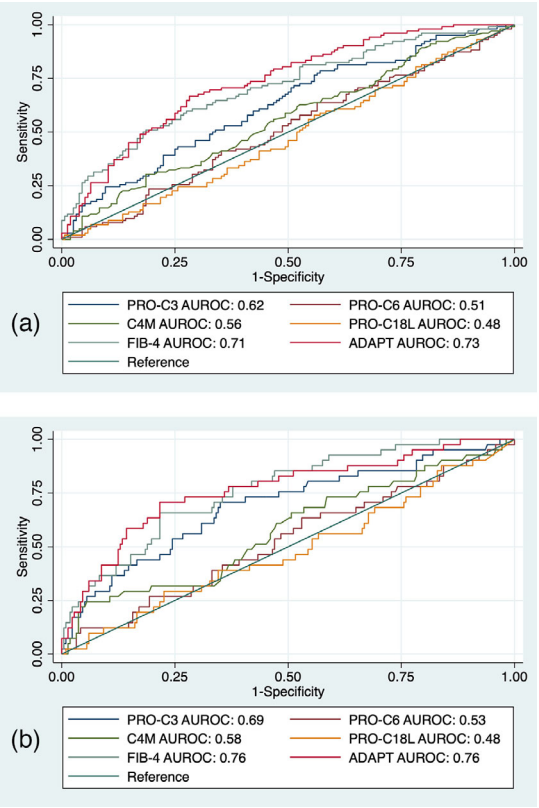

Fig. 1 *Area under the curve (AUC) for PRO-C3, PRO-C6, C4M, PRO-C18L, ADAPT (age, diabetes, PRO-C3, platelet count), and the Fibrosis-4 Index (FIB-4), for detecting patients with a liver stiffness measurement (LSM) >8 kPa (a) and patients with an LSM >12 kPa (b) in patients with metabolic dysfunction-associated steatotic liver disease.*

Table 1. *Baseline characteristics in patients with metabolic dysfunction-associated steatotic liver disease (MASLD) and a liver stiffness measurement (LSM) above, or equal to and below, 8 kPa.*

	LSM ≤8 kPa, $n = 161$ (IQR or %)	LSM >8 kPa, $n = 108$ (IQR or %)	*p*-Value
Male sex	76 (47%)	64 (59%)	0.052
Age (years)	53 (43–62)	59 (54–67)	<0.001
BMI (kg/m^2)	30 (27–34)	32 (29–35)	0.001
Waist circumference (cm)	105 (98–114)	112 (106–120)	<0.001
Waist/Hip ratio	0.99 (0.94–1.04)	1.03 (0.96–1.07)	0.001
Smoking			0.938
Never	86 (53%)	60 (56%)	
Past	62 (39%)	40 (37%)	
Current	13 (8%)	8 (7%)	
Diabetes type 2	54 (34%)	71 (66%)	<0.001
Hypertension	94 (58%)	87 (81%)	<0.001
Hyperlipidemia	132 (82%)	80 (74%)	0.119
AST (U/L)	34 (26–44)	47 (32–65)	<0.001
ALT (U/L)	50 (30–80)	60 (43–88)	0.001
Albumin (g/L)	41 (38–44)	40 (38–43)	0.13
Platelets (×10^9)	239 (205–282)	223 (182–265)	0.017
Creatinine (mmol/L)	72 (60–82)	73 (63–84)	0.28
Fasting glucose (mmol/L)	6.1 (5.7–7.0)	6.9 (6.1–8.2)	<0.001
HbA1c (mmol/mol)	39 (35–44)	46 (38–56)	<0.001
VCTE			
LSM (kPa)	5.8 (5.0–6.7)	10.9 (9.3–14.6)	<0.001
M-probe	95 (59%)	48 (44%)	
XL-probe	61 (38%)	57 (53%)	
CAP (dB/m) ($n = 178$)	314 (280–344)	337 (302–366)	0.002
PRO-C3 (ng/mL)	10.9 (8.8–14.6)	13.3 (10.1–19.0)	0.001
PRO-C6 (ng/mL)	12.1 (9.3–16.8)	12.7 (9.5–17.4)	0.76
C4M (ng/mL)	204 (166–246)	218 (178–259)	0.094
PRO-C18L (ng/mL)	13.0 (9.0–19.4)	12.2 (8.8–18.2)	0.674
FIB-4	1.06 (0.70–1.43)	1.54 (1.03–2.13)	<0.001
ADAPT	5.01 (4.40–5.98)	6.16 (5.27–7.48)	<0.001

Abbreviations: AST, aspartate aminotransferase; ADAPT, (age, diabetes, PRO-C3, platelet count); ALT; alanine aminotransferase; BMI; body mass index; CAP, controlled attenuation parameter; FIB-4, Fibrosis-4 Index; IQR, interquartile range; LSM, liver stiffness measurement; VCTE, vibration-controlled transient elastography.

The PPV for ADAPT versus FIB-4 was 49% versus 46% and the NPV was 86% versus 83%, respectively. Test characteristics for all biomarkers and scores in detecting patients with an LSM >8 kPa are presented in Table 2.

The decision curves showed that ADAPT had a higher net benefit than FIB-4 or VCTE in everyone across a threshold probability range of 15%–40% (Fig. 2a). A probability threshold of 20% would, for example, mean that we are willing to do five VCTE to find one patient with an LSM >8 kPa. In our cohort, this corresponded to a cut-off of 0.5 for FIB-4 and 4.05 for ADAPT (Fig. 3). Using these cut-offs, we would be doing a VCTE in 254 patients using FIB-4 and 241 patients using ADAPT, where they both would identify 106 of 108 patients with an LSM >8 kPa. Translated into a clinical context, this means that ADAPT could reduce the number of unnecessary VCTEs used as a first-line test. The opt-out test trade off at a probability threshold of 20% was 14 tests of ADAPT to avoid one unnecessary VCTE. Comparing ADAPT with FIB-4 at the same threshold probability of 20%, the opt-out test

Table 2. *Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) in detecting patients with a liver stiffness measurement >8 kPa.*

	Cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC (95%CI)
PRO-C3	8.126	90	22	44	76	0.62 (0.55–0.69)
PRO-C6	7.66	90	8	40	52	0.51 (0.43–0.58)
C4M	152	90	19	43	74	0.56 (0.49–0.63)
PRO-C18L	4.888	90	10	40	59	0.48 (0.41–0.55)
FIB-4	0.7352	90	30	46	83	0.71 (0.64–0.77)
ADAPT	4.708	90	37	49	86	0.73 (0.67–0.79)

Abbreviations: ADAPT (age, diabetes, PRO-C3, platelet count); CI, confidence interval; FIB-4, fibrosis-4 index.

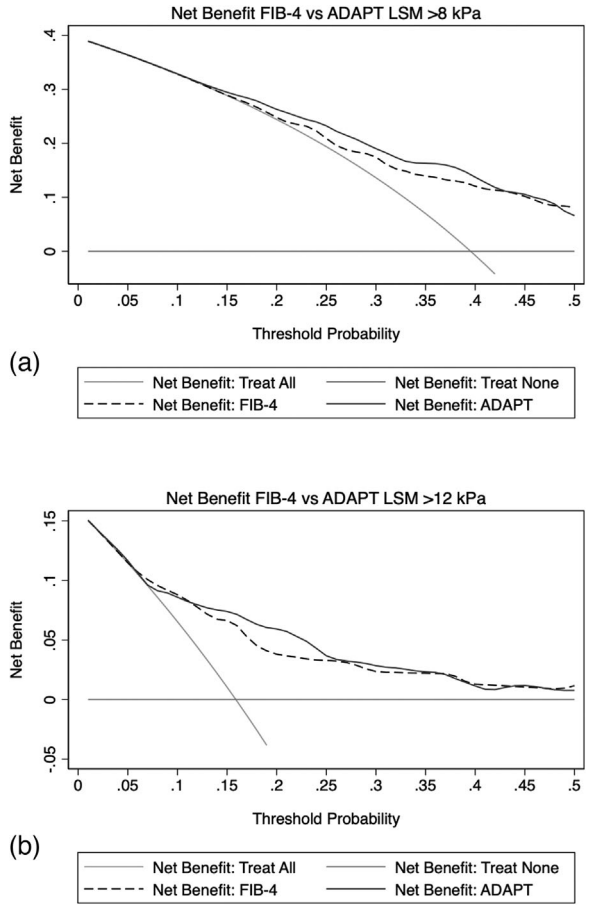

Fig. 2 *Decision curve comparing ADAPT (age, diabetes, PRO-C3, platelet count) and the Fibrosis-4 Index (FIB-4) in detecting patients with a liver stiffness measurement (LSM) >8 kPa (a) and patients with an LSM >12 kPa (b). 'Treat all' corresponds to the strategy of classifying everyone with an LSM >8 kPa (a) or >12 kPa (b) using vibration-controlled transient elastography (VCTE), and 'treat non' corresponds to the strategy of not classifying anyone as having advanced fibrosis. The y-axis displays net benefit, and the x-axis displays threshold probability.*

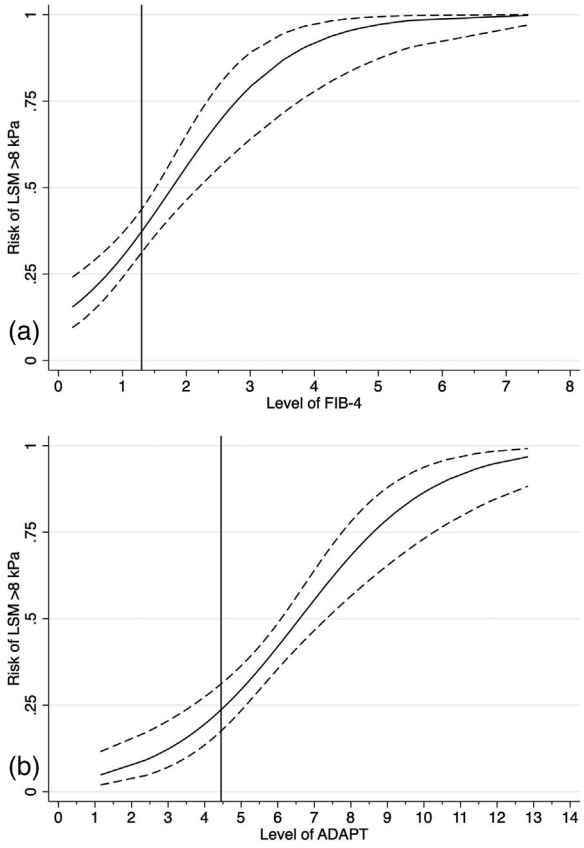

Fig. 3 *Predicted risk of a liver stiffness measurement >8 kPa based on the level of Fibrosis-4 Index (FIB-4) (a) and ADAPT (age, diabetes, PRO-C3, platelet count) (b). Y-axis is the predicted risk (0–1) based on a logistic regression model. X-axis is the level of FIB-4 (a) or ADAPT (b). The dashed lines are the 95% confidence interval. The reference line is for a FIB-4 of 1.3 (a) and an ADAPT of 4.45 (b).*

Table 3. *Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) in detecting patients with a liver stiffness measurement >12 kPa.*

	Cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC (95%CI)
PRO-C3	8.477	90	21	19	91	0.69 (0.60–0.78)
PRO-C6	7.868	90	10	17	82	0.53 (0.43–0.62)
C4M	152	90	17	18	88	0.58 (0.48–0.68)
PRO-C18L	4.788	90	9	17	81	0.48 (0.38–0.58)
FIB-4	1.01	90	41	23	95	0.76 (0.68–0.84)
ADAPT	4.647	90	25	20	93	0.76 (0.68–0.85)

Abbreviations: ADAPT (age, diabetes, PRO-C3, platelet count); CI, confidence interval; FIB-4, fibrosis-4 index.

trade off was 20 tests of ADAPT to avoid one additional unnecessary VCTE compared to 20 tests of FIB-4. Calibration of FIB-4 and ADAPT to detect and LSM >8 kPa is presented in Fig. S2.

Used as a potential second-line test following a FIB-4 >1.30 ($n = 134$), an ADAPT score using the suggested cut-off at 4.45, as an example, had a sensitivity of 100% and a specificity of 11.1% to detect an LSM >8 kPa.

Evaluation of test characteristics at the >12 kPa cut-off

The prevalence of patients with and LSM >12 kPa was 16.7% ($n = 45$), baseline characteristics in patients with an LSM >12 kPa or ≤12 kPa are presented in Table S3.

Among the individual biomarkers, PRO-C3 had the highest discriminatory ability to detect patients with an LSM >12 kPa with an AUC of 0.69 (95%CI = 0.60–0.78) (Fig. 1b). The ADAPT score and FIB-4 had almost equal AUCs, 0.76 (95%CI = 0.68–0.85) and 0.76 (95%CI = 0.68–0.84), respectively ($p = 0.93$) (Fig. 1b). At a set sensitivity of 90%, FIB-4 had the highest specificity of 41%, followed by ADAPT (25%). The PPV and NPV at a sensitivity of 90% were 23% and 95% with FIB-4 and 20% and 93% with ADAPT. Test characteristics for all biomarkers and scores in detecting patients with an LSM >12 kPa are presented in Table 3.

The decision curve analysis showed that ADAPT had a higher net benefit compared to FIB-4 across a threshold probability range of approximately 15%–25%, after which the net benefit was similar between the two tests. Using a probability threshold of 20% as an example, this corresponded to a cut-off of 1.9 for FIB-4 and 6.70 for ADAPT (Figs. 2b and 4). Using these cut-offs, we would be doing

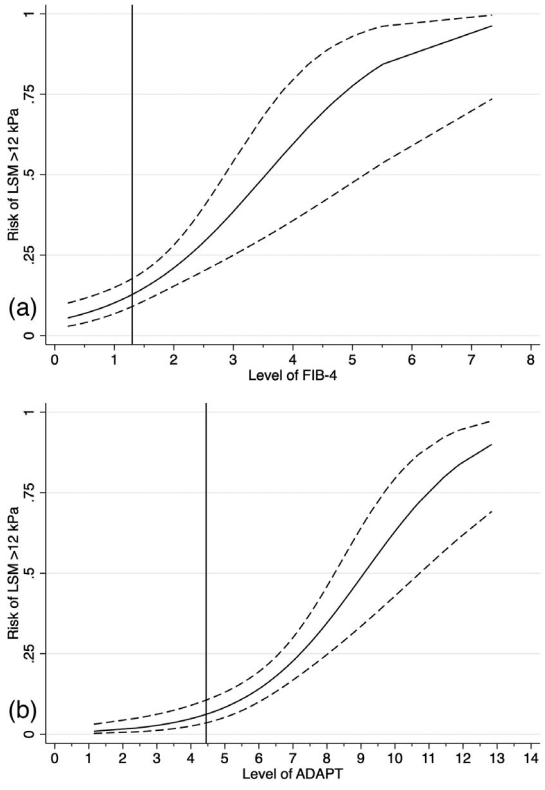

Fig. 4 *Predicted risk of a liver stiffness measurement >12 kPa based on the level of fibrosis-4 index (FIB-4) (a) and ADAPT (age, diabetes, PRO-C3, platelet count) (ADAPT) (b). Y-axis is the predicted risk based on a logistic regression model. X-axis is the level of FIB-4 (a) or ADAPT (b). The dashed lines are the 95% confidence interval. The reference line is for a FIB-4 of 1.3 (a) and an ADAPT of 4.45 (b).*

a VCTE in 63 patients using FIB-4 and 66 using ADAPT, identifying 23 and 27 patients with an LSM >12 kPa, respectively. Translated into a clinical context this means that if we are willing to do an LSM by VCTE in five patients to find one with an

LSM >12 kPa, 2 tests of ADAPT as a first-line test can be used to avoid one unnecessary VCTE. Comparing ADAPT with FIB-4 at the same the threshold probability of 20% the opt-out test trade off was 10 tests of ADAPT to avoid one additional unnecessary VCTE compared to 10 tests of FIB-4. Calibration of FIB-4 and ADAPT to detect and LSM >12 kPa is presented in Fig. S3.

DISCUSSION

In this cross-sectional study of Swedish patients with MASLD, the ADAPT score had the highest discriminatory ability to detect patients with an LSM >8 kPa with an AUC of 0.73. Among the individual collagen cleavage products, PRO-C3 had the highest discriminative ability to detect both patients with an LSM >8 kPa and >12 kPa. For the composite scores, ADAPT had a higher specificity (37%) than FIB-4 (30%) at a fixed sensitivity of 90%. Translated into a clinical context, ADAPT and FIB-4 would identify 9 of 10 patients with an LSM >8 kPa as suggested by guidelines to be a reasonable cut-off for further investigation by hepatologists, but the rate of false positives would be 63% and 70%, respectively. However, using the FIB-4 cut-off of 1.3 in this cohort showed the caveats of this test as a first-line test to exclude patients with an LSM >8 kPa, with a high risk of both false negatives and false positives.

Even though there was no statistically significant difference between the discriminatory ability of ADAPT and FIB-4 in detecting patients with an LSM >8 kPa, the results from the calibration and net benefit analysis suggest that ADAPT may be a better option to use as a first-line test instead of FIB-4. The opt-out test trade off suggested that 20 tests of ADAPT would avoid one additional unnecessary VCTE compared to 20 tests of FIB-4 if they were to be used as first-line tests. This applies if we are willing to do five VCTE to find one patient with an LSM >8 kPa. The optimal threshold probability may differ depending on the clinical settings with varying prevalence of advanced fibrosis in the target population. For example, by using the suggested cut-off of 1.3 for FIB-4 to detect an LSM >8 kPa we indirectly base the number of VCTEs we are willing to do to find one patient using that cut-off. A more reasonable approach would be to decide the number of VCTE we are willing to do based on the clinical setting. The cut-off of the biomarker would then be based on that, and a cut-off used in one situation may not be suitable in another setting. Other aspects, such as cost-effectiveness and availability of noninvasive tests and VCTE, must also be taken into consideration when the clinical utility of a test is evaluated. ADAPT and PRO-C3 are currently not commercially available, and the potential cost of ADAPT and PRO-C3 in commercial use is currently unknown. The cost of a FIB-4 in a Swedish setting is approximately 0.2 euros per FIB-4. If we were to assume that the cost of ADAPT would be 100 times that of FIB-4, the cost to avoid one additional VCTE to find a patient with an LSM >8 kPa, at a probability threshold of 20%, would be 396 euros (the difference in cost between 20 tests of ADAPT and 20 tests of FIB-4).

When investigating the performance of the biomarkers in detecting patients with an LSM >12 kPa, the discriminatory ability of the ADAPT score was equal to FIB-4 (AUC 0.76). The NPV was high for all the evaluated tests and above 90% using the cut-offs for FIB-4 and ADAPT at a fixed sensitivity of 90%, where FIB-4 had a higher specificity (41%) compared to ADAPT (25%). The NPV was >90% for both FIB-4 and ADAPT using the established cut-offs of 1.3 and 4.45, respectively. The ADAPT score was slightly better calibrated than FIB-4 at a lower predicted risk of detecting patients with an LSM >12 kPa. When analyzing the decision curve at a probability threshold of 20% (i.e., the cut-off value where, among those sent to VCTE based on the test, 20% or 1 in 5 will have an LSM >12 kPa), both tests had higher net benefit than the strategy of testing all with VCTE. ADAPT performed better than FIB-4 at this threshold probability, where the opt-out test trade off suggested that 10 tests of ADAPT would avoid one additional unnecessary VCTE compared to 10 tests of FIB-4. At a probability threshold of 10%, there was almost no difference between the ADAPT score and FIB-4. In a clinical setting, it might be appropriate to consider doing more VCTE exams to find patients with an LSM >12 kPa compared to finding patients with an LSM >8 kPa. The added value of using ADAPT instead of FIB-4 to detect patients with an LSM >12 kPa is even more uncertain than in detecting patients with an LSM >8 kPa. FIB-4 performs better to detect patients with an LSM >12 kPa compared to detecting patients with an LSM >8 kPa. This is expected since FIB-4 was developed to detect fibrosis stages 3–4, and not an LSM of >8 kPa [9]. Altogether, these results underline that further studies are needed to delineate how many VCTE exams that are reasonable to perform to find patients with

elevated liver stiffness and higher risk of advanced fibrosis. For the time being, we suggest that FIB-4 should remain the first-line test to first-hand rule out an LSM >12 kPa in patients with MASLD. This may be especially important in low-prevalence settings, where more data on ADAPT as a first-line test is needed.

Eslam et al. investigated the performance of both FIB-4 and ADAPT in 713 patients with MASLD, using LSM with a cut-off of 9.6 kPa to define advanced fibrosis, and found almost identical diagnostic performance [11]. FIB-4 had an AUC of 0.60 and ADAPT of 0.59, respectively, in detecting advanced fibrosis [11]. The prevalence of advanced fibrosis was 2.9%. A direct comparison of the results between this study and the abovementioned study is hard since they used another cut-off of LSM (9.6 kPa) in a primarily Asian population. This emphasizes that the difference in performance between FIB-4 and ADAPT is dependent on both the clinical setting and the prevalence of advanced fibrosis.

A superior discriminatory ability of ADAPT in detecting advanced and significant fibrosis in MASLD has been highlighted in several studies [6, 11, 14–16]. A LITMUS consortium based study by Vali et al. found an AUC of 0.77 compared to 0.73 for FIB-4 in 444 patients with MASLD [14]. A smaller study by Armandi et al. found an AUC of 0.80 for ADAPT compared to 0.72 for FIB-4 in 96 patients [15]. Eslam et al. demonstrated a superior performance of PRO-C3 and ADAPT in two hospital-based cohorts [11]. This was also concluded in a study of 851 biopsy-verified MASLD patients from China, where the prevalence of advanced fibrosis was 12.1%. ADAPT and PRO-C3 were superior in detecting advanced fibrosis compared to FIB-4 [6]. The AUC of ADAPT was 0.87, whereas FIB-4 had an AUC of 0.75 [6]. Most data in these studies stem from tertiary centers with a high pretest probability of advanced fibrosis but did not report net benefit. They indicate a better performance of algorithms including PRO-C3 in populations with earlier stages of fibrosis and in populations with a relatively high prevalence of fibrosis, even higher than in our cohort. Hence, using ADAPT as a first-line test in the general population may lead to an overutilization of LSM by VCTE as second-line testing.

Even though the individual collagen neo-epitopes evaluated in this study did not perform as well as FIB-4 or ADAPT, a future role of these biomarkers as diagnostic or prognostic tools cannot be ruled out. Since they mirror the formation and degradation of collagens associated with liver fibrosis [17], they might not correlate to the present liver stiffness, being less applicable in cross-sectional studies such as this. Future studies are needed to evaluate how longitudinal changes in these biomarkers are associated with disease progression or regression, and also as treatment response biomarkers. Further research is also needed regarding ADAPT as a second-line test, although our data suggest a low specificity at only 11% for the currently suggested cut-off at 4.45.

The strengths of our study include a large sample size, including patients with MASLD from various regions of Sweden. The distribution of men and women was even, and the proportion of patients with an LSM >12 kPa was relatively low, indicating a high external validity. We used the LSM cut-offs suggested in the American Gastroenterological Association (AGA) clinical practice update [4]. Other cut-offs have also been suggested but have not been investigated in the current study. There were several limitations: First, VCTE was only performed at one time point, with a risk of false positives. Second, patients were included at tertiary centers, possibly explaining the high prevalence of an LSM >8 kPa. The decision curve analysis is based on the prevalence of an LSM >8 kPa and >12 kPa in this cohort and should be interpreted with caution. Our primary analysis included patients with VCTE, partly since biopsy was not available in the full cohort. However, the main reason was that our aim was to evaluate the clinical pathway of first-line testing with noninvasive tests and their performance in detecting patients with an LSM >8 kPa, as a VCTE is the recommended second-line test in guidelines. The aim was not to identify patients with fibrosis stage 3 or 4 on liver biopsy. Other liver fibrosis biomarkers, such as ELF, which is recommended as an alternative to FIB-4 as a first-line test in the AGA clinical practice update [4], were not evaluated. This should be performed in future studies. Lastly, few patients had a high predicted probability, making calibration estimates uncertain in that range.

Conclusion

The PRO-C3-based ADAPT algorithm seems to be marginally better than FIB-4 in detecting patients with an LSM >8 kPa. FIB-4 performed similar to

ADAPT in detecting patients with an LSM >12 kPa. However, the added clinical value of ADAPT instead of FIB-4 as a first-line test in patients with MASLD is uncertain. The advantage of ADAPT compared to FIB-4 tends to depend on the prevalence of advanced fibrosis and the number of VCTE we are willing to perform to find one patient with elevated liver stiffness. For now, we suggest that FIB-4 remain the first-line noninvasive test, with the primary goal to rule out patients with LSM >12 kPa. However, we also conclude that better first-line tests are needed, where a possible utility of ADAPT might be further evaluated.

Author contributions

Study conception and design: Hannes Hagström and Mattias Ekstedt. *Acquisition of data*: Hannes Hegmar, Thomas Wiggers, Patrik Nasr, Johan Vessby, Stergios Kechagias, Hanns-Ulrich Marschall, Nils Nyhlin, Åsa Danielsson Borssén, Mattias Ekstedt, and Hannes Hagström. *Statistical analysis*: Hannes Hegmar. *Analysis and interpretation of data*: Hannes Hegmar, Hannes Hagström, Rickard Strandberg, and Diana Julie Leeming. *Drafting of manuscript*: Hannes Hegmar. *Critical revision*: All

Conflict of interest

The authors declare no conflicts of interest.

Funding information

HHa was supported by grants from The Swedish Research Council, The Swedish Cancer Society, and Region Stockholm. HHe has received supporting grants from Region Stockholm.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Guarantor of the article

All authors approved the final version of the article, including the authorship list.

References

1 Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE, et al. The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol.* 2022;**7**(9):851–61.
2 Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol.* 2023;**78**(6):1966–86.
3 Taylor RS, Taylor RJ, Bayliss S, Hagström H, Nasr P, Schattenberg JM, et al. Association between fibrosis stage and outcomes of patients with non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Gastroenterology.* 2020;**158**(6):1611–1625.e12.
4 Wattacheril JJ, Abdelmalek MF, Lim JK, Sanyal AJ. AGA clinical practice update on the role of noninvasive biomarkers in the evaluation and management of non-alcoholic fatty liver disease: expert review. *Gastroenterology.* 2023;**165**(4):1080–88.
5 Udompap P, Therneau TM, Canning RE, Benson JT, Allen AM. Performance of American Gastroenterological Association Clinical Care Pathway for the risk stratification of patients with non-alcoholic fatty liver disease in the US population. *Hepatology.* 2023;**77**(3):931–41.
6 Tang LJ, Ma HL, Eslam M, Wong GL, Zhu PW, Chen SD, et al. Among simple non-invasive scores, Pro-C3 and ADAPT best exclude advanced fibrosis in Asian patients with MAFLD. *Metabolism.* 2022;**128**:154958.
7 Daniels SJ, Leeming DJ, Eslam M, Hashem AM, Nielsen MJ, Krag A, et al. ADAPT: an algorithm incorporating PRO-C3 accurately identifies patients with NAFLD and advanced fibrosis. *Hepatology.* 2019;**69**(3):1075–86.
8 Boursier J, Vergniol J, Guillet A, Hiriart JB, Lannes A, Le Bail B, et al. Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. *J Hepatol.* 2016;**65**(3):570–78.
9 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**(6):1317–25.
10 Nielsen MJ, Nedergaard AF, Sun S, Veidal SS, Larsen L, Zheng Q, et al. The neo-epitope specific PRO-C3 ELISA measures true formation of type III collagen associated with liver and muscle parameters. *Am J Transl Res.* 2013;**5**(3):303–15.
11 Eslam M, Wong GL, Hashem AM, Chan HL, Nielsen MJ, Leeming DJ, et al. A sequential algorithm combining ADAPT and liver stiffness can stage metabolic-associated fatty liver disease in hospital-based and primary care patients. *Am J Gastroenterol.* 2021;**116**(5):984–93.
12 Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* 2016;**352**:i6.
13 Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Eur Urol.* 2018;**74**(6):796–804.
14 Vali Y, Lee J, Boursier J, Petta S, Wonders K, Tiniakos D, et al. Biomarkers for staging fibrosis and non-alcoholic steatohepatitis in non-alcoholic fatty liver disease (the LITMUS project): a comparative diagnostic accuracy study. *Lancet Gastroenterol Hepatol.* 2023;**8**(8):714–25.
15 Armandi A, Rosso C, Younes R, Leeming DJ, Karsdal MA, Caviglia GP, et al. Cross-sectional and longitudinal performance of non-invasive tests of liver fibrosis in patients with non-alcoholic fatty liver disease. *J Clin Med.* 2023;**12**(2):650.

16 Nielsen MJ, Leeming DJ, Goodman Z, Friedman S, Frederiksen P, Rasmussen DGK, et al. Comparison of ADAPT, FIB-4 and APRI as non-invasive predictors of liver fibrosis and NASH within the CENTAUR screening population. *J Hepatol.* 2021;**75**(6):1292–300.
17 Karsdal MA, Nielsen SH, Leeming DJ, Langholm LL, Nielsen MJ, Manon-Jensen T, et al. The good and the bad collagens of fibrosis – their role in signaling and organ function. *Adv Drug Deliv Rev.* 2017;**121**:43–56.

Correspondence: Hannes Hegmar, C1:77, Division of Hepatology, Karolinska University Hospital, 141 86 Stockholm, Sweden.
Email: hannes.hegmar@ki.se)

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supplementary Figure S1: Flow chart of the inclusion and exclusion of patients.

Supplementary Figure S2: Calibration curves of FIB-4 (A) and ADAPT (B) in detecting patients with a liver stiffness measurement (LSM) >8 kPa. Y-axis is observed risk of an LSM > 8 kPa. X-axis is predicted risk of an LSM >8 kPa. The diagonal dashed line indicates agreement between predicted and observed risks. The light blue area is the 95% confidence interval. The red horizontal line above the x-axis displays observed risk among patients with an LSM >8 kPa (=1), and an LSM equal or below 8 kPa (=0).

Supplementary Figure S3: Calibration curves of FIB-4 (A) and ADAPT (B) in detecting patients with a liver stiffness measurement (LSM) >12 kPa. Y-axis is observed risk of an LSM >12 kPa. X-axis is predicted risk of an LSM >12kPa. The diagonal dashed line indicates agreement between predicted and observed risks. The light blue area is the 95% confidence interval. The red horizontal line above the x-axis displays observed risk among patients with an LSM >12 kPa (=1), and an LSM equal or below 12 kPa (=0).

Supplementary Table S1: Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), and AUC in detecting patients with an LSM >8 kPa using previously proposed cut-offs of FIB-4 and ADAPT.

Supplementary Table S2: Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), and AUC in detecting patients with an LSM >12 kPa using previously proposed cut-offs of FIB-4 and ADAPT.

Supplementary Table S3: Baseline characteristics in patients with metabolic dysfunction-associated steatotic liver disease (MASLD) and a liver stiffness measurement (LSM) above, or equal and below 12 kPa. ■

