Biomarker Qualification Letter of Intent (LOI) Content Elements

Administrative Information

1. Submission Title:

"Circulating Biomarkers for Diagnosis of Non-Alcoholic Steatohepatitis (NASH)"

DDTBMQ000084

2. Requesting Organization: Foundation for the National Institutes of Health Biomarkers Consortium
11400 Rockville Pike, Suite 600
North Bethesda, MD 20852
Phone: (301) 402-5311
Website: https://fnih.org/what-we-do/biomarkers-consortium

• Primary Point of Contact:

Tania Kamphaus, PhD
Scientific Program Manager
Phone: (301) 435-6247
Email: TKamphaus@fnih.org

• Alternate Point of Contact:

Arun J. Sanyal, MD
Professor
Virginia Commonwealth University School of Medicine
1250 E Marshall Street, Richmond, VA 23298
Phone: (804) 827-0716
Email: arun.sanyal@vcuhealth.org

• Supporting or participating organizations or individuals

FNIH Biomarkers Consortium | **N**on **I**nvasive Bio**M**arkers of Meta**B**olic **L**iver Disease**E** (NIMBLE)

Project Steering Committee: Arun Sanyal, MD (VCU); Rohit Loomba (UCSD); Claude Sirlin, MD (UCSD); Anthony Samir, MD (MGH-Harvard); Sudha Shankar, MD (AZ); Roberto Calle, MD (Pfizer); Sarah Sherlock, PhD (Pfizer)

3. Submission Dates:

LOI submission date (original)- **02/26/2019**

LOI submission date (updated)- **10/15/2019**

Drug Development Need Statement

Nonalcoholic fatty liver disease (NAFLD) affects 30% of adults in the United States and is characterized by accumulation of excess triglycerides in the liver. By definition, it is not due to consumption of harmful amounts of alcohol. It has two principal subtypes: (1) nonalcoholic fatty liver (NAFL) and (2) nonalcoholic steatohepatitis (NASH). NASH, especially in those who have developed some fibrosis, progresses to cirrhosis and end-stage liver disease, much more frequently than NAFL. The identification of patients at risk of progression to cirrhosis and clinically meaningful adverse outcomes, and the development of effective therapies for these patients, is a public health priority. Both NAFL and NASH are however relatively asymptomatic until advanced disease is present and cannot be readily diagnosed or distinguished from each other by clinical and routine laboratory tests. The diagnosis and staging of NASH therefore relies on a liver biopsy with histological assessment of the liver; and such an assessment represents the reference standard for this purpose. Liver biopsies are however invasive, painful and carry a limited but real risk for catastrophic complications including death. Furthermore, histologic evaluation using liver biopsy is significantly limited by moderate accuracy at best, due to the small amount of material gathered for examination relative to the size of the liver, and due to interpreter variability.

Akin to the situation in clinical practice described above where diagnosis of NASH is dependent on histologic assessment, at present, the only way to identify patients with NASH for participation in clinical trials is a liver biopsy. Patients are selected to undergo screening with liver biopsy through an assessment of their probability of having NASH that depends on clinical history and non-specific laboratory values. Unfortunately, selection of subjects for biopsy via clinical means is imperfect and leads to a high frequency of biopsies in patients that end up not meeting the histopathologic entry criteria for NASH necessary for inclusion in clinical trials. The screen-failure rate across phase 2b-3 trials for NASH range from 50-80%. *The lack of availability of reliable non-invasive means of enriching the population of subjects undergoing liver-biopsy that are likely to meet the histopathological criteria for NASH with fibrosis represents a major barrier for identification of subjects eligible for pharmacological intervention and therefore for enrollment in clinical trials. The identification of such markers would enable clinical trials and accelerate the development or pharmacologic therapeutic interventions for NASH. These are patients that on biopsy show evidence of active NASH (defined by presence of steatohepatitis and a NAFLD activity score (NAS) of 4 or more) with fibrosis score ≥2 and are thus identified as a population of patients with NASH with a higher probability to progress to adverse liver outcomes due to the stage of fibrosis.*

This underscores the urgent need for robust, reliable, readily deployable non-invasive tools that identify subjects with a high probability of meeting the histopathologic criteria for active NASH with fibrosis stage 2 or more. In considering those with NASH and fibrosis stage 2 or higher, it is also important to distinguish those with cirrhosis (Stage 4 fibrosis) from those with earlier stage disease because the disease biology and clinical course are different requiring different approaches to therapy and assessment. In this LOI, we focus on those with NASH with high disease activity (NAS ≥ 4) and fibrosis stages 2-3 that we refer to as "at risk" NASH. A separate LOI will be submitted for non-invasive tools to diagnose NASH with cirrhosis.

The drug development need that will be addressed by this qualification is the diagnostic enrichment of "at-risk NASH" patients for inclusion in drug development trials.

The development of NASH follows a complex interplay of mechanisms where genetic susceptibility coupled with environmental factors such as sedentary lifestyle and excess energy intake results in hepatic steatosis, insulin resistance, enhanced di novo lipogenesis, lipotoxicity, hepatocyte oxidative stress and mitochondria disfunction, and activation of inflammatory pathways with inflammatory infiltration, and ultimately resulting, in some patients, to fibrosis and organ damage, including cirrhosis. The multifaceted nature of the pathogenesis of NASH means that the influence of multiple contributory factors needs to be taken into account and quantified if we hope to determine non-invasively the level of

risk of a patient to progress to adverse liver outcomes. This is true whether we are trying to draw a correlation of biomarkers directly with these clinical outcomes, or through and intermediate step such as showing correlation with histological liver findings that are themselves associated with liver outcomes (e.g., fibrosis). A number of circulating and imaging-based biomarkers have been shown to have a degree of correlation with liver biopsy histological assessment. These biomarkers provide read-outs on some of these varied pathways contributing to NASH pathogenesis. However, these biomarkers have not been systematically and independently evaluated. The NIMBLE Project's overarching objective is, through a systematic evaluation of these promising biomarkers (both individually and in a number of potential combinations) to identify a biomarker (or biomarker panel) that will be both robust and sensitive to identify "at-risk" NASH patients.

This application is limited to testing four selected circulating biomarker panels which may eventually be considered as stand-alone panels or in combination to identify active NASH with stage 2-3 fibrosis. The FNIH NIMBLE team selected these biomarkers based on publicly available data (peer-reviewed publication) supporting their analytical and clinical performance, as well as data summaries provided by the biomarker developers. The supporting information for each biomarker is available within the reference section. Summary data for each of the biomarkers is provided in sections below.

The team is aware of several other biomarkers modalities including imaging (ultrasound and MR-based) and functional biomarkers that have sufficient supporting data to warrant additional investigation. **These markers are not part of the current LOI. However, a number of promising imaging markers are being assessed in a parallel study and will be the subject of a separate LOI. A separate LOI will also be submitted for the noninvasive diagnosis of cirrhosis due to NASH.** It is possible that some biomarkers are superior in performance to others. **Decisions on the final biomarker or biomarker panel (to be part of a future qualification plan submission) will be made after all initially selected biomarkers are tested and analyzed.**

The circulating biomarkers that are part of this LOI are listed in the table below. These biomarkers reflect the activity of different pathways contributing to the etiopathogenesis of NASH. It is possible that a combination of these markers more accurately predicts who is the "at risk" NASH patient and correlates better with histological assessment.

Biomarker	**NASH** (inflammation)	**NAS** (composite histological score)	**Fibrosis**	**Composite** (Biomarker with more than one analyte)
NIS 4	**X**		**X**	**X**
OWL Liver	**X**	**X**		**X***
PRO-C3			**X**	**X**
Siemens (ELF)			**X**	**X**

* *Multiple lipid species*

Context of Use Statement (500 characters)

The selected biomarker "panel" (e.g., NIS-4, OWLiver®, ProC3/C6, ELF, or combination thereof) is a diagnostic enrichment biomarker intended for use, in conjunction with clinical factors, to identify patients likely to have liver histopathologic findings of nonalcoholic steatohepatitis (NASH) and with a nonalcoholic fatty liver disease activity score (NAS) ≥4 and liver fibrosis stages 2 or 3 (by Brunt/Kleiner scale); and thus appropriate for inclusion in liver biopsy-based NASH drug development clinical trials focused on pre-cirrhotic stages of NASH.

In the remaining document, we provide information for each one of the four biomarkers to be tested as part of this study.

1. NIS4 Marker

(Please note: information below summarizes information provided by Genfit)

Biomarker Information and Interpretation

1. **Biomarker name: Non-Invasive Score 4 [NIS4]** (microRNA 34a-5p [miRNA 34a], alpha2-macroglobulin [A2M], Hemoglobin A1c [HbA1c], Chitinase-3-Like Protein 1 [CHI3L1, aka YKL40]). Biomarker type: Molecular. Unique molecular ID: **UniProt** (http://uniprot.org/) - alpha2-macroglobulin (P01023), hemoglobin A1c (P69905), CHI3L1/YKL40 (P36222). **miRBase** (http://www.mirbase.org/) - microRNA 34a-5p (miRBase MIMAT000025).

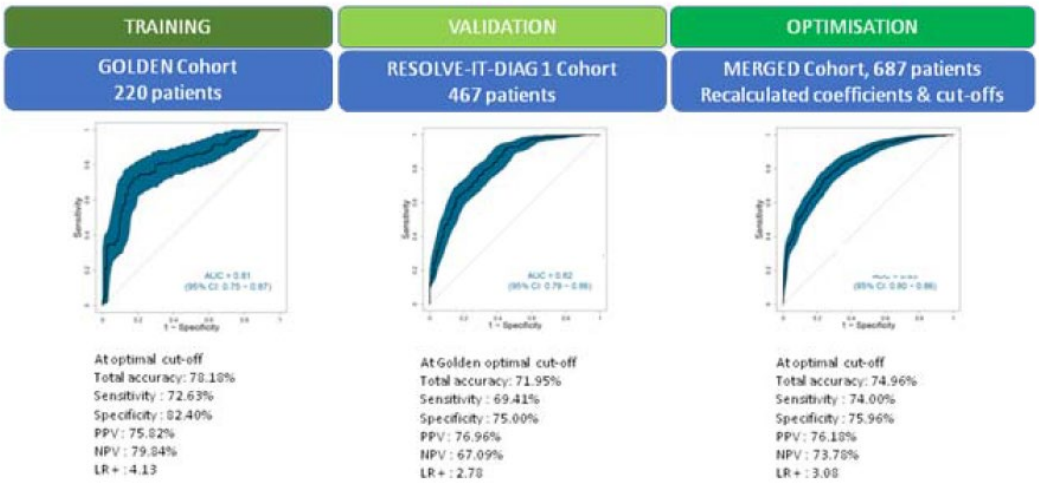
2. **Analytical methods:** The NIS4 NASH Diagnostic Test is an In Vitro Diagnostic Multivariate Assay (IVDMIA) Test that calculates a score based on the quantification of four individual analytes:

a. miR-34a-5p: measured by real time quantitative polymerase chain reaction using a Bio-Rad CFX96 thermo-cycler
b. A2M (Alpha2-Macroglobuline): Measured by immunonephelometry using a Siemens BNII or BNProspect instrument with kits OSAM09 or OSAM15
c. CHI3L1 (Chitinase-3-Like Protein 1): Measured using solid-phase immunoassay with a kit from R&D Systems
d. HbA1c: Measured by HPLC using a Bio-Rad HPLC D10 with kit 220-0101 or Variant II with kit 270-2455

3. Measurement units and limit(s) of detection: The NIS4 NASH Diagnostic Test is an In Vitro Diagnostic Multivariate Assay (IVDMIA) Test that calculates a score based on the quantification of four individual analytes: miR-34a-5p, CHI3L1 (Chitinase-3-Like Protein 1), and A2M (Alpha2-Macroglobuline) and HbA1c. This score has a value ranging 0 to 1 and assigns the risk probability for having active NASH (NAS≥4) with significant fibrosis (F≥2).

4. **Biomarker interpretation and utility**: Genfit NASH NIS4 was developed/trained in the GOLDEN Elafibranor Phase II drug trial and internally validated in a first set of 467 patients screened for RESOLVE-IT Elafibranor Phase III drug trial. As shown in Figure 2, NIS4 demonstrated similar performances with an AUC of 0.82 and 0.81 in the Training and Validation cohorts.

Figure 1: Training, Optimization and Validation of NIS 4.



Analytical Considerations

Analyte	Abbreviation	FDA Status	Measurement	Kit / Instrument	Sample Type
microRNA 34a-5p	miR-34a	RUO	Real time quantitative polymerase chain reaction	Instr: Bio-Rad CFX96 thermo-cycler	Serum; SST Tube
α-2 Macroglobulin	A2M	Cleared	Immuno-Nephelometry	Instr: Siemens BNII or BNProspect Kit: OSAM09 or OSAM15	Serum; SST Tube
Chitinase-3-Like Protein 1	CHI3L1	RUO	Solid-phase Immunoassay	R&D Systems	Serum; SST Tube
Hemoglobin A1c	HbA1c	Cleared	HPLC	Bio-Rad HPLC D-10 with Kit 220-0101 or Variant II with kit 270-2455	Whole blood; EDTA Tube

Table 1: The analytical methods for each measured component of the NIS-4 multi-component biomarker (Data courtesy GENFIT) ***Analytical considerations for individual components of the NIS4 diagnostic test can be provided upon request.***

Clinical Considerations

A tool that can enrich the selection of patients submitted to a liver biopsy, limiting it to subjects with a higher probability of meeting the required histopathologic criteria would be a welcomed tool for both patients and Sponsors. A schematic description of the proposed use of this biomarker is illustrated below in Figure 2.

The biomarker would be used in an adult population (aged 18 or above) considered to have a high likelihood for NASH (e.g., absence of a history of alcohol abuse, clinical characteristics of metabolic syndrome, elevated liver fat content by imaging, and/or elevated liver enzymes) with sufficiently advanced fibrosis to be at high risk of progression to adverse liver outcomes. NIS4 is to be used with a single, high-value cut-off correlating to a diagnostic PPV>80% for efficient referral and high diagnostic confidence. NIS4 diagnostic performance was determined in RESOLVE-IT DIAG. Diagnostic cut-off is set at 0.617 to diagnose patients with active NASH and significant fibrosis (NAS≥4 & F≥2) with a specificity goal of 85% which should yield a high (> 80%) positive predictive value. Sensitivity, specificity, and PPV can be altered by different prevalence in the condition to diagnose, but also in disease distribution in the group above NAS≥4 & F≥2. In addition, the overall performance characteristics of NIS4 would indicate a limited number of false-positive and hence limit the number of liver biopsies on patients that do not meet the histological entry criteria.

Figure 2: NIS 4 Diagnostic Performance.

Cut-Off 0.617	RESOLVE-IT-DIAG cohort
n	487
Prevalence	55%
AUROC	0.82
Accuracy	75%
Sensitivity	60%
Specificity	85%

If an individual yields a NIS4 score below the high cut-off, these patients are recommended for follow-up and re-test at regular intervals at the discretion of the attending physician. Given the slow progression of NASH in most patients, initially missing a subject with F≥2 should not result in adverse outcomes so long as continued monitoring is maintained. Results from RESOLVE-IT longitudinal samples and biopsies should support the determination of the ideal frequency of re-testing as a function of a baseline or initial NIS4 value. The NIMBLE study will help provide such additional, external information.

Supporting Information

Support for the role of miR-34a

Due to their role in modulating both cellular processes and expression levels in disease processes, circulating miRNAs have been proposed as biomarkers for different pathologies, including liver disease, cardiac diseases, diabetes, cancer, kidney alterations, neurodegenerative diseases, allergy and asthma, amongst many others.

By using samples collected during Elafibranor clinical trials (GOLDEN-505 and RESOLVE-IT) and other external cohorts, GENFIT identified >10 miRNAs whose circulating levels are altered in patients with active NASH and significant fibrosis. Among the ones identified, miR-34a demonstrated the highest diagnostic value in identifying patients with active NASH and significant fibrosis. Circulating levels of miR-34a increase with both NASH disease grade (e.g. NAS) and fibrosis stage. Furthermore, miR-34a outperformed all other markers previously reported in the literature (notably CK18-M30, CK18-M65, PIIINP) for detection of active NASH and significant fibrosis. In line with its potential roles in NASH pathophysiology and arguing in favor of its use for diagnosis of active NASH with significant fibrosis, miR-34a is present in very low levels in presumed "healthy" blood donors (Table 2).

Table 2: Expression levels of miR-34a in healthy vs diseased individuals (*data courtesy Genfit*)

"Healthy" donors Mean (sd) (nbr)	NTBT GOLDEN Mean (sd) (nbr)	TBT GOLDEN Mean (sd) (nbr)	p-value NTBT vs TBT
40 (6) (100)	218 (15) (110)	435 (34) (160)	<0.0001

NTBT = Not=To-Be-Treated patients; TBT = To-Be-Treated patients; GOLDEN = GOLDEN Phase II Elafibranor drug trial; sc = standard deviation; nbr = numbers

Support for the role of A2M

Alpha-2-macroglobulin or A2M is the second most important biomarker in the NIS4 test. A2M is a plasma inflammation glycoprotein with anti-protease and anti-fibrinolytic activities. It is produced in the liver and its levels are greatly increased when fibrosis develops. A2M has an important role in protease inhibition thus allowing an increase in collagen production by reduction of collagenase activity.

Support for the role of CHI3L1/YKL-40

Chitinase-3-like protein (also known as YKL-40) is a secreted glycoprotein implicated in liver fibrosis. CHI3L1 is expressed by many different types of cells, such as monocytes, macrophages, endothelial cells, and chondrocytes. It has also been shown that CHI3L1 is able to suppress matrix metalloproteinase production, such as MMP-1, MMP-3 and MMP-13, induced by IL-1 and TNF-α. Evidence suggests that CHI3L1 may utilize its chitin binding ability to communicate with other signal transduction pathways to modulate various physiologic processes, such as inflammation, apoptosis, tissue remodeling, cell growth, and angiogenesis.

Those performances have been presented in various posters at AASLD 2016/2017, EASL 2017/2018 and NASH Biomarkers 2017 (Reference A: NIS4 References #'s 10-13).

LABCORP has signed a license agreement to perform NIS4 for clinical research worldwide, which includes performing the 4 assays and calculating the score with the required assays techniques and instruments, and within a CLIA/CAP certified environment. The scoring information can be requested from Genfit and LABCORP.

2. OWLiver®

(Please note: information below summarizes information provided by One Way Lipidomics)

Biomarker Information and Interpretation

1**. Biomarker name: OWLiver®** (triglycerides panel with varying saturation of associated fatty acids plus Body Mass Index [BMI]); Biomarker type: Molecular. Unique molecular ID: ID for OWLiver is not formally stablished because it is an "in vitro diagnosis" based in an algorithm combining 25 different molecular (trigylicerides) markers.

2. **Analytical methods:** OWLiver® is a lipidomic serum test based on two sequential analyses for discrimination between normal liver and NAFLD and then, NASH from NAFL (simple steatosis). This is based on two logistic regression algorithms

based on a panel of 11 and 20 triglycerides, respectively, including BMI in the classification algorithm (Patent: Priority# EP16382320.6). It is performed in an LC-MS system and indicated for adults with BMI > 25 suspected of having NASH.

For the purpose of this LOI, we will be focusing on OWLiver's ability to distinguish NASH from NAFL.

The diagnostic method measures two panels of triglycerides in serum and perform a combined analysis with the relative intensities of those triglycerides and the body mass index (BMI) of the patients to obtain the probability of the patients to present NAFL, NASH or no NAFLD. Data are processed using the TargetLynx application manager for MassLynx software (Waters Corp.). A set of predefined retention time mass-to-charge ratio pairs, corresponding to metabolites included in the analysis, are fed into the program. Associated extracted ion chromatograms are then peak detected and noise reduced in both the LC and MS domains such that only true metabolite-related features are processed by the software. Normalization factors are calculated for each metabolite as described by Martínez-Arranz et al. 2015. Logistic regression analysis is applied to achieve a predictive signature capable of first discriminating between NAFLD and no NAFLD, and second between NASH and NAFL, introducing BMI as a continuous variable (Martínez-Arranz et al., 2015).

Receiver operating characteristic (ROC) curve analysis is used to assess the discriminatory power. Overall diagnostic accuracy for a given two-class comparison was done by the area under the ROC curve (AUROC) with its associated standard error. Sensitivity, specificity, and positive and negative predictive values (PPV and NPV, respectively) are estimated. The full description of the current methods is described in Barr *et al.,* 2012; Martínez-Arranz *et al*., 2015; and in Mayo *et al.,* 2018.

3. **Measurement units and limit(s) of detection:**

Sensitivity of methodology

Sensitivity for the test that discriminates between NAFL and NASH: 0.70

Specificity of methodology

Specificity for the test that discriminates between NAFL and NASH: 0.81

PPV and NPV where available

PPV for the test that discriminates between NAFL and NASH: 0.81

NPV for the test that discriminates between NAFL and NASH: 0.69

4. **Biomarker interpretation and utility (*as provided by OWL*):** The diagnostic performances of the validated tests show an area under the receiver operating characteristic curve, sensitivity, and specificity of 0.79 ± 0.04, 0.70, and 0.81, respectively, for the discrimination between NASH and NAFL in a population of 192 biopsied adult patients with biopsy proven NAFLD (7 normal, 109 NAFL and 76 NASH) from five hospitals in Spain and the Czech Republic, age 18-70, and no known chronic disease aside from obesity or type 2 diabetes (Mayo *et al.,* 2018). The effect of glucose level was not taken into account in the algorithms. When the algorithm is applied to populations with type 2 diabetes, an increase in the error rate was observed for glucose levels >136 mg/dL (Bril *et al.,* 2018; Mayo *et al.,* 2018). When the analysis was performed excluding patients with glucose levels >136 mg/dL, the area under the receiver operating characteristic curve for the discrimination between NASH and NAFL increased to 0.81 ±0.04 with sensitivity and specificity of 0.73 and 0.80, respectively. (Mayo *et al.,* 2018).

The test does not take into account the potential role of diabetes (hyperglycemia) or ethnicity. The test performance decreases when applied to a multiethnic cohort of patients with type 2 diabetes, although the performance increases if a specific subset of patients is selected to mirror the characteristics of the population used for the development of the test (Bril, *et al*. 2018).

Lower Limit of Quantification (LLOQ) of methodology/assay

The test is based on a semi-quantitative metabolomics methodology. The process includes an intra- and inter-batch normalization process based on multiple internal standards and pool calibration samples approach (Martinez-Arranz *et al.,* 2015).

Details of any algorithms employed for data analysis

The patented algorithms are detailed in the patent "DIAGNOSTIC METHODS BASED ON LIPID PROFILES". Application for PCT in Europe and USA PCT/EP2017/066915. International publication number: WO2018/007511

Once the relative intensities of the analytes are normalized, the data obtained is introduced in the algorithm. Logistic regression analysis was applied to this set of normalized relative intensities to distinguish between NASH and NAFL. The objective of logistic regression analysis is to find the best fitting model to explain the connection between the dichotomous variable and a set of independent conditions. In this case, those independent conditions were the lipidomic features of the panel of selected TGs and the BMI, which is introduced in the algorithms as a continuous variable.

Firstly, the algorithm is applied to distinguish between normal liver (No NAFLD) vs. NAFLD individuals. Secondly, to those previously classified as NAFLD, the algorithm is applied to discriminate between NAFL and NASH. The probabilities of both tests will be reported together with the final diagnostic result.

Analytical Considerations (*as provided by OWL*):

Analytical variability of methodology (CV)

Average of the %CV of the triglycerides included in the two algorithms:
%CV intra-batch = 3.98%
%CV inter-batch = 5.94%
%CV inter-laboratory = 6.02%

Test-retest metrics

Algorithm for the discrimination between simple steatosis and NASH:
Performance in the discovery cohort: AUROC = 0.95; Se = 0.83; Sp = 0.94; PPV = 0.89; NPV = 0.90.
Performance in an independent blinded cohort: AUROC = 0.79

Sensitivity and specificity of methodology

Algorithm for the discrimination between simple steatosis and NASH:
Sensitivity = 0.83.
Specificity = 0.94.

PPV and NPV

Algorithm for the discrimination between simple steatosis and NASH: PPV = 0.89; NPV = 0.90.

Analytical variability of methodology (CV)

Two different serum extracts (two commercial sera) are used as quality controls (QC) to assess the data quality. These QCs are used to verify that the assay has been correctly done. The first one (QC_Calibration) is used for both internal standard intrabatch and interbatch correction, whereas the second one (QC_Validation) is used to evaluate the overall reproducibility and system precision and also to validate the algorithm result. The results of the QC samples provides the

basis of accepting or rejecting the run data. Each run contains at least 3 QC_Calibration serum and at least 3 QC_Validation serum samples, that are evenly distributed through the run.

Acceptance criteria for QC samples for assay qualification: The precision (%CV) calculated on the relative intensities of the TG from the n determinations must be less than or equal to 20%, for both QC_Calibration and QC_Validation. The algorithm result for the QC_Validation should be equal to the previously established value at the beginning of the study.

Preferred matrix for measurement: Human serum, obtained from blood collected under fasting conditions (≥8 h) and extracted into vacutainer SST II Advance tubes or similar, without anticoagulant and with a separating gel (Vacutainer System Blood Set needles, or similar).

Sample processing procedures

Triglyceride profiles of the serum samples are obtained as follows: serum samples are mixed with sodium chloride (50 mM) and chloroform/methanol (2:1) in 1.5 mL micro-tubes on ice. The extraction solvent is spiked with a TG compound, TG (13:0/13:0/13:0), not detected in unspiked human serum extracts. TGs are glycerides in which the glycerol is esterified with three fatty acid groups. TG (13:0/13:0/13:0) in particular, consists of one chain of tridecanoic acid at the C-1 position, one chain of tridecanoic acid at the C-2 position and one chain of tridecanoic acid at the C-3 position of the glycerol group.After brief vortex mixing, the samples are incubated for 1 hour at -20 °C. Once centrifuged at 16,000g for 15 minutes, the organic phase is collected, and the solvent removed. The dried extracts are then reconstituted in acetonitrile/isopropanol (1:1), centrifuged (16,000g for 5 minutes), and transferred to vials for ultra-high-performance LC coupled to MS time of flight (UHPLC-MS-TOF) analysis.

Stability of analytes over time

Long term -80 °C storage stability is at least 2 years.

Clinical Considerations

The OWLiver® test was performed in an original cohort of 467 NAFLD patients (90 normal liver, 246 steatosis and 131 NASH) that was subsequently validated in a separate cohort of 192 patients (7 normal liver, 109 steatosis, 76 NASH). Further clinical validation has been achieved in a study performed together with University of Florida (Gainsville, Fl, USA) in a subgroup of Caucasian individuals with well controlled diabetics.

Supporting Information

Developed in an ultra-high performance-LC coupled to a Synapt G2 QTOF mass spectrometer from Waters Corp. (Milford, USA) and subsequently implemented in a Xevo G2 QTOF system from Waters Corporation. OWLiver tests are performed in Spain under a Reference Laboratory Certificate of the Department of Health, Vasque Government, Spain. OWLiver is certified as meeting ISO 9001:2015 and ISO 13485:2016. PROC3-6: Procollagen III fragment are measured by ELISA at Nordic Bioscience; CAP accredited and ISO9001 certified. ||

3. PROC3

(Please note: information below summarizes information provided by Nordic BioSciences)

Biomarker Information and Interpretation

1. **Biomarker Name:** Pro-Collagen 3 [**PRO-C3** fragment]. Biomarker type: Molecular. **UniProt** (http://uniprot.org/) - PRO-C3 (P02461) cleavage fragments.
2. **Analytical methods:** PRO-C3 has been shown to be highly associated with liver fibrosis across a broad range of liver fibrosis etiologies and is a promising non-invasive biomarker. A new non-invasive score system for the identification

of patients with advanced liver fibrosis was created by combining a marker of fibrogenesis, PRO-C3, with various risk factors for advanced liver fibrosis (age, diabetic status and platelet count). The derived score was termed ADAPT after the score components[6]. The score was developed from a cohort of 150 patients from Australia and was further validated in a multi-center international cohort of 281 patients.

$$ADAPT = exp\left(log_{10}\left(\frac{Age \times PRO\text{–}C3}{\sqrt{Platelets}}\right)\right) + Diabetes$$

Units of measurement, key cut-off values, "normal vs abnormal" in absolute and relative units

Cut-off values are only available for PRO-C3 and are shown in Table 3 for the different diseases. For the other markers, we have determined a "normal range" value, based on a healthy population.

Cut-offs for PRO-C3 in different liver diseases. Non-alcoholic fatty liver disease, NAFLD; Non-alcoholic steatohepatitis, NASH; ALD, Alcoholic liver disease; HCV, Hepatitic C virus; HBV, Hepatitis B Virus, Sensitivity; Sens, Spec; Specificity, PPV; Positive Predictive Value, NPV; Negative Predictive Value, AUC; Area Under the Curve. Advanced fibrosis defined as stages 3-4 (bridging fibrosis or cirrhosis)

For the purposes of this COU, we are evaluating the role of this biomarker in diagnosing fibrosis.

Table 3: Units of measurement, key cut-off values, "normal vs abnormal" in absolute and relative units

Etiology	AUC	Cutoff value (ng/mL)	Prevalence (%)	Sens/Spec (%)	PPV/NPV (%)
NASH vs NAFL	0.78	14.71[6]	42.5	75.3/71.5	66.1/79.7
NAFLD (advanced fibrosis)	0.83	15.6[6]	23.1	86.2/70.4	46.7/94.4
HCV (advanced fibrosis)	0.71	20.2[7]	50.2	44.6 / 87.2	77.8 / 61.0
HBV (advanced fibrosis)	0.71	19.2	29.6	45.3 / 85.1	74.1 / 62.3
ALD (advanced fibrosis)	0.88	20.6	24.3	55.8 / 94.8	86.6 / 78.0

3. **Measurement units and limit(s) of detection:**

Lower Limit of Quantification (LLOQ) of methodology/assay

See Table 3.

Analytical variability of methodology (CV)

See Table 3.

Test-retest metrics

Approval of the unknown samples x = the mean concentration of two or three replicate measures of an unknown sample.

LLOQ ≤ x < ULOQ

For unknown samples within the measuring range with a CV ≤ 15%, the mean concentration is approved and reported as the result. Unknown samples with a CV > 15% at the first analytical run (R1) must be reanalyzed. If the sample is applied

from the second analytical run (R2), this result is reported. If not, the sample must be re-analyzed a third time. If the sample is approved from the third analytical run (R3), this result is reported. If the sample is not approved in neither R1, R2 nor R3, the result of R3 is reported with the comment "Inaccurate test result" and the result is noted as a deviation in the analytical report. If the sample volume is insufficient to run R1, R2 or R3, the sample is reported as Not Done (ND) with the comment "Insufficient sample amount".

x > ULOQ

Unknown samples above ULOQ, or above STD A if STD A < ULOQ, should be re-analyzed at a higher dilution than the one used in R1. If the sample is approved from R2, this result is multiplied by the dilution factor and reported. If the sample volume is insufficient to run R2, the sample is reported as ND with the comment "Insufficient sample amount".

x < LLOQ

For unknown samples below LLOQ, the LLOQ value (corrected for dilution) of the specific assay is reported as the result with the comment "Below lower limit of quantification". For samples below LLOQ, the %CV is not considered. x = "RANGE?"

Unknown samples reported as "RANGE?" are identified as being below LLOQ or above ULOQ based on the optical density value and handled as described under 3. x < LLOQ or 2. x > ULOQ, respectively. 5. Sensitivity of methodology N/A

PPV and NPV where available- Only available for PRO-C3, see table 2 for PPV and NPV values for the detection of advanced fibrosis.

4. **Biomarker interpretation and utility (*provided by Nordic BioSciences*):** Collagens are key players in fibrogenesis and fibrosis resolution and may provide distinct information regarding the nature of fibrosis depending on their localization and function in the extracellular matrix (ECM). The ECM is subdivided into the basement membrane and the interstitial matrix, each containing specific types of collagens. By assessing specific fragments of collagens, it may be possible to separate tissue formation from tissue degradation, i.e. measurement of pro-peptides and matrix metalloproteinase (MMP) degraded collagen fragments (Pais *et al*., 2013). Nordic Biosciences has developed and validated a panel of serological biomarkers, all of which target specific fragments of a given protein including both formation and degradation markers of the same protein. We hypothesize that markers of collagen formation, fibrogenesis, are decreased as a function of anti-fibrotic therapy, whereas markers of collagen degradation, fibrosis resolution, will be increased. Therefore, the purpose is to evaluate the ability of specific collagen formation and degradation markers to identify "at risk" NASH patients and responders of therapy using serum samples from the NIMBLE trial.

Biological rationale to distinguish NAFL from NASH: NAFLD is an umbrella term encompasses a range of phenotypes, common amongst them is the excessive accumulation of fat within the hepatocytes of the liver. NASH is the progressive manifestation of NAFLD and is associated with a poorer prognosis and faster progression of disease (Pais *et al.,* 2013; McPherson *et al.,* 2015; Singh *et al.,* 2015). In addition to hepatic steatosis, a NASH diagnosis is confirmed by the presence of hepatocyte ballooning and lobular inflammation (Chalasani *et al*., 2018). This metabolic hepatic inflammatory milieu is the principal driver of disease progression (Schauppan *et al*., 2018). Hence, NASH is considered the progressive form of NAFLD and that patients with NASH have a greater disease activity (Brunt *et al*., 2011). PRO-C3 is a marker of true type III collagen formation, it specifically targets the site at which the N-terminal propeptide is cleaved off by ADAMTS27. As such as well as being a marker of liver fibrosis, PRO-C3 has been demonstrated to be a marker of fibrogenesis and disease activity. PRO-C3 has shown to be highly associated with the NAFLD activity score (NAS) and can significantly discriminate

between patients with simple steatosis and NASH. (Daniels *et al.,*2019; Luo *et al.,* 2018; Leeming *et al. J. Hepatol,* 2017; Leeming *et al. NASH Biomarkers Work, 2017)*.

Nordic Bioscience has developed a range of ELISA based assays that can evaluate the remodeling of the extracellular matrix of the liver and the development of hepatic fibrosis. These assays specifically detect the MMP-mediated degradation and formation of type I. II, III, IV, V and VI collagen. ***For the purposes of this Letter of Intent the key marker of interest is PRO-C3. Additional markers, in particular PRO-C6, will be analyzed to assess their contribution to the information provided by the measurement of PRO-C3 in isolation.*** The analytical performance of the PRO-C3 and PRO-C6 fragments are provided in Table 4. PRO-C3 is a marker of true type-III collagen formation and has been shown to be highly related to liver fibrosis. This marker is CE marked as a serological marker.

TABLE 4: NORDIC BIOSCIENCE BIOMARKERS *(Data courtesy Nordic Biosciences)*

Assay	Target	LLOQ (ng/mL)	Detection Range (ng/mL)	CV	Intraassay variation (%)	Interassay Variation (%)	Matrices	Volume required (µL)
PRO-C3[1]	N-terminal pro-peptide of type III collagen	2.6	2.6 – 116.0	≤10	1.5 - 8.3	3.6 - 13.3	Serum & plasma	50
PRO-C6[17]	C-terminal pro-peptide of type VI collagen	0.8	0.8 – 134.0	≤10	1.1 - 5.3	3.4 - 12.4	Serum & plasma	50

Analytical Considerations

Preferred matrix for measurement

Serum

Cross-validation across matrices

Serum, EDTA plasma, Heparin plasma and citrate plasma can all be used for assessment of the markers.

Minimum and maximum matrix volumes required for measurement (eg: in duplicate Minimum volume 50 µL per marker. 400 µL in total.

Detailed sample collection procedures

Screening and on study samples for laboratory assessments shall be collected after an overnight fast (8 hours, water is acceptable). Collect whole blood into serum tubes.

Detailed sample processing procedures

Blood is collected into plain tubes and left at room temperature for 30 min to cloth, then centrifuged at 5000 rpm for 10 minutes. All liquid is transferred to new tubes and again centrifuged at 5000 rpm for 10 min, where after serum is transferred to clean tube and stored at -80°C for subsequent analyses of serological biomarkers.

Sample instructions

The samples stored at -80oC, should be shipped on dry ice to our facility in Denmark (Nordic Bioscience A/S, Herlev Hovedgade 205-207, 2730 Herlev, Denmark). Samples should be stored at -80°C until analysis. The assays have been validated up to four freeze-thaw cycles. The assay is accepted with a recovery of 100±20%. Analyte stability of serum samples have been tested at 1 hour, 2 hours, 24 hours, 72 hours and one week at 4°C, 20°C and 37°C. The assay is accepted with a recovery of 100±20%.

Clinical Considerations

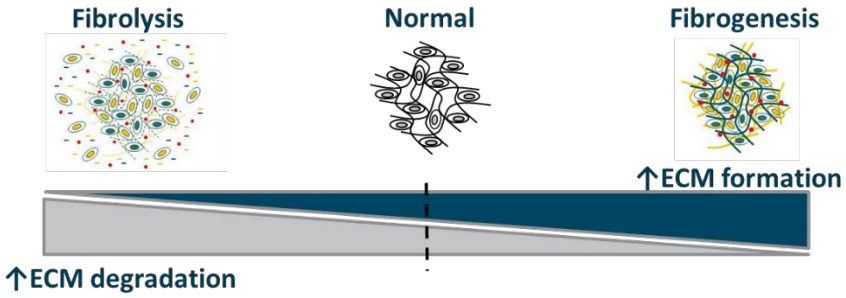
According to the recently published draft guidance document by the FDA, patients with a NASH activity score (NAS) ≥ 4 with at least 1 point each in inflammation and ballooning along with a NASH CRN fibrosis score greater than stage 1 fibrosis but less than stage 4 fibrosis should be considered eligible for clinical trials investigating non-cirrhotic NASH12. PRO-C3 is intended to identify individuals with NAFLD at high risk of having high NASH disease activity (i.e. NAS ≥ 4, based on the NASH-CRN scoring system). Potential candidates are pre-screened according to set eligibility criteria based on clinical factors consistent with a risk for NASH. PRO-C3 will be assessed in serum samples from individuals who meet the other eligibility criteria. If the PRO-C3 levels are above a selected cut-off, individuals are deemed to be at high risk of having high NASH disease activity and fibrosis stages 2 or higher, and will be recommended to undergo a confirmatory liver biopsy. The use of PRO-C3 is hypothesized to reduce the number of biopsies required in patient screening. ***Reference values will be generated from patients enrolled in the NIMBLE study, along with other relevant NAFLD/NASH cohorts as well as published reference values in the literature.***

The incorporation of PRO-C3 into the screening strategy for entry into clinical trials regarding NASH will decrease the number of biopsies for future clinical trials. As such, PRO-C3 will allow for a more efficient and less burdensome development plan. At this point in the development no risks have been identified.

Supporting Information:

There is an unmet need for high quality biomarkers that can safely and reproducibly predict the stage of the disease progression for patients with Chronic Liver Disease (CLD). The requirement for such markers has intensified due to the high prevalence globally of non-alcoholic fatty liver disease (NAFLD), including NASH. The ECM of the liver is a dynamic structure and undergoes continuous remodeling. In this process, old tissue is degraded, and new tissue is formed in a delicate equilibrium. Disruption of this balance may lead to conditions such as fibrosis, which is defined by the accumulation of ECM, as illustrated in Figure 3.

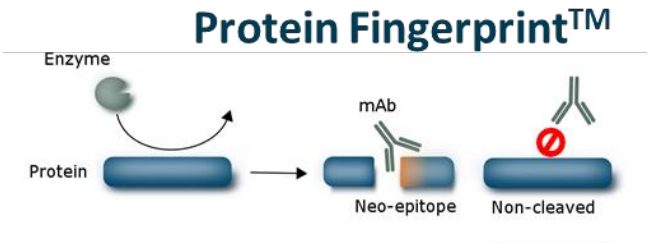
Figure 3: Balance of ECM turnover



Nordic Bioscience has developed a plethora of blood-based biomarkers for the assessment of ECM remodeling that utilize neo-epitope technology (figure 4). Liver fibrosis is characterized by imbalanced ECM remodeling. This process results generation of degradation and formation products of extracellular and intracellular proteins generated by disease-specific proteases expressed at the site of injury and released into circulation. The unique combination of a protein and its protease constitute a protein fingerprint of the ongoing changes in a specific tissue and may provide information about the pathogenesis of disease and serve as biomarker target. The neo-epitope approach has been successfully utilized in other disease fields such as rheumatology. As the neo-epitope biomarkers represent the turnover products of the fibrotic structure this type of markers may consequently be related to diagnosis, prognosis, and efficacy of intervention. The Protein Fingerprint technology further enables measurement of sub-pools of the same protein, as these fragments may provide different information emphasizing the importance of distinguishing each sub-pool from others. This method has proven to be more sensitive and more accurate than more routinely used diagnostic and prognostic tools. In most fibroproliferative diseases several different pathways often contribute to the development of the disease. Therefore, quantification of a single factor in a highly complex process may not reflect the nature of the disease due to redundancy. Serological neo-epitope biomarkers represent the end-product of tissue destruction and may more accurately quantify the effects of converging pathways.

We envision using these unique protein fingerprints markers as single determinants and in a multi-marker approach. The information from several unique markers combined in such a multi-marker approach may increase the diagnostic precision and accuracy when describing NASH patients. The biomarkers developed by Nordic Bioscience can be measured in serum and plasma by specific ELISAs utilizing neo-epitope specific monoclonal antibodies (mAbs).

Figure 4: Neo-epitope biomarkers detecting ECM remodeling. Formation of detectable neo-epitopes generated by cleavage of ECM proteins by specific proteases, and released into the bloodstream. In the absence of protease cleavage, the ECM protein is not recognized by the antibody.



4. ELF Marker

(Please note: information below summarizes information provided by Siemens)

Biomarker Information and Interpretation

Biomarker name: Enhanced Liver Fibrosis [ELF] panel (Procollagen III n-terminal peptide [PIIINP], hyaluronic acid [HA], tissue inhibitor of metalloproteinase 1 [TIMP-1]); Biomarker type: Molecular. Unique molecular ID: **UniProt** (http://uniprot.org/) - procollagen III N-terminal propeptide (P02461), hyaluronic acid (Q92839), TIMP-1 (P01033).
The ADVIA Centaur® systems Enhanced Liver Fibrosis (ELF) test is an *in vitro* diagnostic multivariate index assay intended to provide a single ELF score by combining in an algorithm the quantitative measurements of hyaluronic acid (HA), amino-terminal propeptide of type III procollagen (PIIINP) and tissue inhibitor of metalloproteinase 1 (TIMP-1) in human serum using the ADVIA Centaur systems.

2. **Analytical methods:** The ELF Score is an In Vitro Diagnostic Multivariate Assay (IVDMIA) Test that calculates a score based on the quantification of three individual analytes: PIIINP, HA and TIMP-1 measured through automated through an ADVIA Centaur or ADVIA Centaur XP system.
To calculate the ELF score manually for the ADVIA Centaur and ADVIA Centaur XP systems, first obtain results for the ADVIA Centaur HA, PIIINP, and TIMP-1 assays, and then use the following equation to calculate the ELF score:

- ELF score = 2.278 + 0.851 ln(CHA) + 0.751 ln(CPIIINP) + 0.394 ln(CTIMP-1)
- Concentrations (C) of each of the constituents are in ng/mL.

Note: The ELF score is a unitless numerical value.

Units of measurement, key cut-off values, “normal vs abnormal” in absolute and relative units. Note the ELF score is a unitless numerical value.

Lower Limit of Quantification (LLOQ) of methodology/assay

N/A in IFU. LoD and LoB are provided below.

- HA
  - Limit of Detection (LoD) = 1.6 ng/mL,
  - Limit of Blank (LoB) = 0.8 ng/mL
- PIINP
  - Limit of Detection (LoD) = 0.5 ng/mL,
  - Limit of Blank (LoB) = 0.1 ng/mL.

- TIMP-1
    - Limit of Detection (LoD) = 3.5 ng/mL
    - Limit of Blank (LoB) = 0.1 ng/mL.

Sensitivity of methodology / Specificity of methodology

For the lower cutoff:

Area Under the Receiver Operating Characteristic Plot (AUROC) = 0.786 (95% CI: 0.755 to 0.817)

- Sensitivity = 88.6% (319/360, 95% CI: 84.9 to 91.7%)
- Specificity = 34.6% (194/561, 95% CI: 30.6 to 38.7%)

For the upper cutoff:

AUROC = 0.859 (95% CI: 0.826 to 0.893)

- Sensitivity = 65.4% (106/162, 95% CI: 57.6 to 72.7%)
- Specificity = 89.7% (681/759, 95% CI: 87.3 to 91.8%)

4. **Biomarker interpretation and utility *(all data provided by Siemens)***
Interpretation of the CE-marked ELF* score is as follows:

Table 5: Interpretation of the CE-marked ELF score

ELF Score	Severity of Liver Fibrosis
< 7.7	None to mild
≥ 7.7– < 9.8	Moderate
≥ 9.8	Severe

The ADVIA Centaur ELF kit is not FDA approved and not commercially available in US. It is CE-marked and available outside the US. A Research Use Only (RUO) ELF testing service is exclusively available from Siemens Healthcare Laboratory, Berkeley CA.

Analytical Considerations

Analytical variability of methodology (CV)

Precision was evaluated according to the CLSI protocol EP5-A2.13 Samples were assayed in 3 replicates 2 times a day for 20 days (n = 120 replicates per sample). Each of the component assays were calibrated according to their individual calibration intervals.

The following results were obtained from testing performed on one ADVIA Centaur XP system using one reagent lot and one calibrator lot:

Table 6: Analytical variability of methodology (CV)

Mean ELF Score	Within-Run SD	Between-Run SD	Total SD
6.98	0.07	0.04	0.11
7.12	0.04	0.03	0.08
8.95	0.03	0.04	0.09
11.05	0.03	0.04	0.08
14.51	0.04	0.03	0.08

3. Clinical variability of methodology in populations and conditions where measured

HA Precision Data

Precision was evaluated according to CLSI guideline EP5-A2.14 Samples were assayed in 3 replicates 2 times a day for 20 days (n = 120). The following results were obtained from testing performed on 1 ADVIA Centaur XP system using 1 reagent lot and 1 calibrator lot. The instrument was calibrated on the first run of day 1 and recalibrated after 14 days.

Table 7: HA Precision Data

Mean (ng/mL)	Within-Run %CV	Between-Run %CV	Total %CV
10.9	5.2	2.4	5.9
12.6	4.2	4.3	6.1
19.7	4.5	3.8	7.5
51.4	3.6	5.3	7.7
60.4	4.2	0.0	5.5
211.5	3.9	3.2	6.6
315.8	3.8	3.3	5.7
966.4	4.5	3.2	7.7

PIIINP Precision Data: Precision was evaluated according to CLSI guideline EP5-A2.14 Samples were assayed in 3 replicates 2 times a day for 20 days (n = 120). The following results were obtained from testing performed on 1 ADVIA Centaur XP system using 1 reagent lot and 1 calibrator lot. The instrument was calibrated on the first run of day 1 and recalibrated after 28 days.

Mean (ng/mL)	Within-Run %CV	Between-Run %CV	Total %CV
1.9	5.0	3.4	6.6
3.0	2.3	2.1	3.4
4.3	1.7	3.1	3.6
4.9	3.3	5.4	6.8
8.1	1.9	1.6	2.9
10.8	2.2	3.4	4.4
36.3	1.8	2.0	2.8
104.3	1.9	3.1	3.7
124.1	2.2	2.8	3.5

Table 8: PIIINP Precision Data

TIMP-1 Precision Data: Precision was evaluated according to CLSI guideline EP5-A2.15 Samples were assayed in 3 replicates 2 times a day for 20 days (n = 120). The following results were obtained from testing performed on 1 ADVIA Centaur XP system using 1 reagent lot and 1 calibrator lot. The instrument was calibrated on the first run of day 1 and recalibrated after 28 days.

Table 9: TIMP-1 Precision Data

Mean (ng/mL)	Within-Run % CV	Between-Run % CV	Total % CV
73.3	2.4	3.2	4.0
81.2	2.5	4.4	5.1
142.2	1.8	2.7	3.3
253.5	1.9	5.7	6.0
333.6	1.6	2.7	3.1
524.1	1.8	4.9	5.2
718.3	1.8	3.1	3.6
1050.6	2.0	3.0	3.6

Diagnostic accuracy of the ELF® score to identify significant fibrosis and clinical events during follow-up has been shown to be high (Puigvehõ M*, et al. 2016)*

Test-retest metrics.

Samples may be retested, and patient sample dilutions performed in accord with the instructions for use recommendations.

Preferred matrix for measurement: Serum (SST or Red Top) tubes are used for sample collection – transfer into 2.0 mL SC Micro Tube Sarstedt 72.694.406.

Cross-validation across matrices: Serum only. Cross-validation not needed.

Detailed sample collection procedures:

Serum is the only recommended sample type for these assays. The following recommendations for handling and processing blood samples are provided by the Clinical and Laboratory Standards Institute (CLSI):

- Collect all blood samples observing universal precautions for venipuncture. Handle all samples as if capable of transmitting disease.
- Allow samples to clot adequately before centrifugation.
- Keep tubes stoppered at all times.
- Test samples as soon as possible after collecting.
- Do not use specimens with obvious microbial contamination.

Clinical Considerations

The clinical utility of ELF has been investigated in a range of CLD, including NAFLD, ALD, and Viral Hepatitis. Studies have explored the usefulness of ELF in both primary care and specialty settings. Advantages to ELF include the ability to collect a serum sample at any routine phlebotomy site, and the stability of samples (properly stored) for transport and analysis if testing is not available on-site. Outside the US, ELF has been clinically validated for both diagnostic and prognostic applications. Studies support the utility of ELF vs. biopsy, as well as compared to non-invasive technologies such as imaging for liver elasticity. ELF allows the categorization of patients for inclusion/exclusion of severe fibrosis with a high degree of accuracy and thus, in conjunction with clinical characteristics, and possibly in combination with other markers, either molecular or imaging in nature, may enrich biopsy results for patients that have a higher likelihood of meeting the criteria for inclusion in clinical trials (fibrosis stage ≥ 2 in histological assessment).

A growing number of studies support the utility of ELF in aiding assessment of NAFLD/NASH patients for significant fibrosis.

•ELF performs well for the detection of advanced fibrosis in NAFLD patients (both adults and children), including patients with early signs of steatohepatitis. (Reference D: ELF References #'s 1, 4, 8-12, 14, 16, 25, 28, 34, 37-43). All recent liver

guidelines (EASL-EASD-EASO, AASLD, BSG, NICE) are in agreement in recommending screening for advanced fibrosis (histological stage Kleiner F3-4) in patients diagnosed with NAFLD.

- ELF was ranked number 1 based on both performance and cost estimate analysis in the NICE Guidelines for NAFLD (NICE guideline NG49, 2019).
- ELF is recommended for second line testing in the BSG Guidelines for abnormal LFT's to identify advanced fibrosis (following FIB-4 or NFS) (Newsome *et al.,* 2018)
- ELF (in combination with FIB-4 as the front-line test) significantly increased detection of advanced fibrosis/cirrhosis in NAFLD patients tested in a primary care setting and significantly reduced unnecessary referrals. (Srivastava *et al.,* 2019; Armstrong and Marchesini, 2019)
- Changes in ELF have been correlated to NASH therapy (in development) (Harrison *et al.,* 2019; Harrison *et al.,* 2017; Loomba *et al.,* 2018) Early data suggests a reduction in ELF score is associated with of other parameters of liver fibrosis improvement

The clinical utility of ELF has been investigated in a range of CLD, including NAFLD, ALD, and Viral (Reference D: ELF References #'s 1-30). Studies have explored the usefulness of ELF in both primary care and specialty settings. Advantages to ELF include the ability to collect a serum sample at any routine phlebotomy site, and the stability of samples (properly stored) for transport and analysis if testing is not available on-site site (Reference D: ELF References #'s 8, 18, 20, 31-35). Outside the US, ELF has been clinically validated for both diagnostic and prognostic applications. Studies support the utility of ELF vs. biopsy, as well as compared to non-invasive technologies such as imaging for liver elasticity. ELF allows the categorization of patients for inclusion/exclusion of severe fibrosis with a high degree of accuracy and thus, in conjunction with clinical characteristics, and possibly in combination with other markers, either molecular or imaging in nature, may enrich biopsy results for patients that have a higher likelihood of meeting the criteria for inclusion in clinical trials (fibrosis stage ≥ 2 in histological assessment).

Supporting Information

The Enhanced Liver Fibrosis (ELF) test is a noninvasive blood (serum) lab test designed to assess levels of three major components directly involved in liver matrix metabolism: hyaluronic acid (HA), procollagen III amino-terminal peptide (PIIINP), and tissue inhibitor of matrix metalloproteinase 1 (TIMP-1). The analytes are automatically measured, and the software calculates and reports a unitless numeric score. Increasing ELF scores quantitatively reflect ECM turnover and associated HSC activation, and are linked to both biopsy-proven fibrosis and to prognosis for clinically significant outcomes (1) The ELF score has been well validated against biopsy-proven fibrosis across a range of chronic liver diseases (CLD) in both adult and pediatric populations (Day *et al.,* 2019; Rosenberg *et al.,* 2004; Fagan *et al.,* 2015; Nobili *et al.,* 2009). While differences can exist with form of CLD in terms of how/pattern of the liver is damage, fibrosis progression/regression/repair is a key element of pathology across almost all forms of CLD (Marcellin *et al.,*2018; Bataller *et al.,* 2005; Angulo *et al.,*2015; Dulai *et al.,* 2010).

Common etiologies include nonalcoholic fatty liver disease (NAFLD), alcoholic liver disease (ALD), and viral hepatitis (VH), all of which have robust ELF data associated with assessment of fibrosis severity, including abundant data on the performance vs. biopsy. Changes in the ELF score have been correlated to a change in staging with biopsy (Day *et al.,* 2019; Irvine *et al.,* 2016).

As disease progression is strongly linked to liver fibrosis which in turn reflects ECM activity and subsequent liver scarring, ELF component analytes are released into the blood as a consequence and so likely reflect "real-time" activity. This differs from both biopsy and imaging which are essentially a "look-back" at damage done. Data suggests ECM measurement may therefore be more predictive of future events vs. biopsy. ELF also differs from indirect markers which reflect inflammation and compromised liver function but not directly the ECM activity. ECM measurement includes outcome data from CLD patients where ELF values of >9.8 outperformed biopsy for disease progression. While liver fibrosis and ECM activity is biochemically complex, it is orchestrated primarily by activated hepatic stellate cells (HSCs). Activated HSCs produce

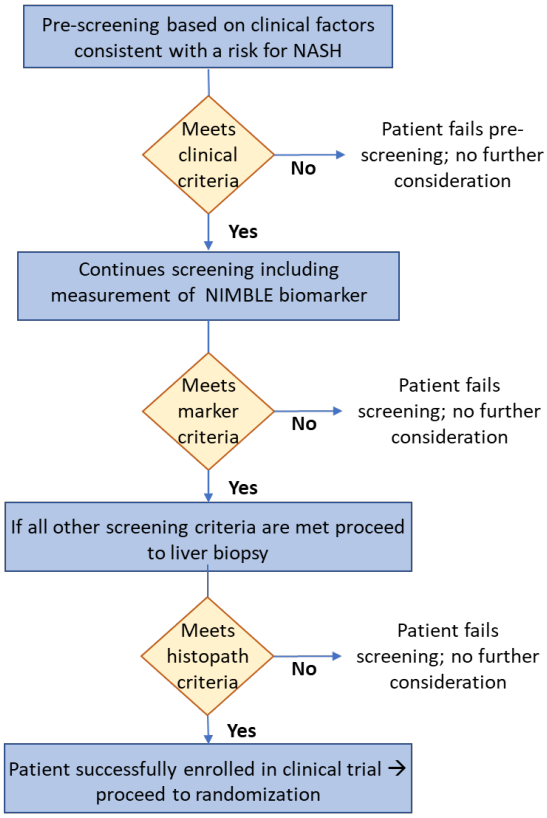
components of the extracellular matrix (ECM), including the analytes measured with the ELF test. The ECM proteins include fibronectin, laminin, collagens, hyaluronic acid (HA), and proteoglycans. Collagen types I, III, IV, and V are prominently expressed within the liver. N-terminal propeptide of procollagen type III (PIIINP) is generated during the synthesis of type III collagen and is involved in fibrin formation. Increased concentrations of PIIINP are seen in liver fibrosis associated with CLD and ECM turnover. HA is an essential component of the ECM and is produced primarily by HSC. As ECM activity and deposition progressively replaces the normal liver parenchyma, producing damage and scar tissue and ultimately disrupting hepatic architecture and function. During this process, many circulating ECM proteins (or derivatives) are released into the blood, including the 3 ELF markers.

Fibrosis of the liver is a largely bidirectional process (Bataller *et al.,* 2005; Ramachandran *et al.,* 2012) and ECM activity can reflect both deposition and repair as fibrosis and repair mechanisms have been linked to ECM-related pathways. Regression and repair are associated in part with upregulation of matrix metalloproteinases (MMPs), which are a family of zinc-dependent endopeptidases capable of degrading ECM deposition and so important to healing. Levels of MMPs are subject to inhibition by tissue inhibitors of metalloproteinases (TIMPs), a family of at least four proteins (TIMP 1–4) which bind MMPs. TIMP-1 overexpression hinders degradation and clearance of the fibrotic matrix, leading to increased levels of interstitial ECM and progressive fibrosis. Additionally, low levels of TIMP-1 may promote hepatic stellate cell apoptosis. Therefore, increased levels of TIMP-1 are associated with inhibition of repair and increased fibrogenesis. By testing for direct markers associated with both ECM deposition and repair, the ELF test provides a direct measure for the assessment of fibrotic activity at the biochemical level.

The three direct markers of the ELF test provide complementary information, and the combined score outperforms both the individual markers and simple scores such as APRI or FIB-4. The performance of the ELF test for liver fibrosis has been well established in the scientific literature and guidelines, and ease of testing and interpretation support routine clinical use and an alternative to invasive biopsy.

Decision Tree

The following decision tree is suggested for the implementation of any of the biomarkers (or combination thereof) described in this LOI, as a diagnostic enrichment biomarker implemented in a biopsy-based clinical trial



Previous Qualification Interactions and Other Approvals (if applicable)

DDTBMQ000084 - Letter of Intent to Qualify Circulating Biomarkers for Diagnosing Non-Alcoholic Steatohepatitis (NASH)

(NIMBLE Project) original submission date : 02/26/2019