Letter of Intent

Letter of Intent from LITMUS for the Biomarker Qualification Program

Table of Contents

1. List of Abbreviations

AASLD	American Association for the Study of Liver Diseases
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BQP	Biomarker Qualification Program
CAP	Controlled Attenuation Parameter
CE	Manufacturer's declaration that the product meets the applicable EC directives
CLD	Chronic liver disease
CLIA	Clinical Laboratory Improvement Amendments
COU	Context of Use
EASL	European Association for the Study of the Liver
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EMA	European Medicines Agency
EU	European Union
FDA	U.S. Food and Drug Administration
HSC	hepatic stellate cells
IDE	Investigational Device Exemption
IFCC	International Federation of Clinical Chemistry
IFU	Information for use
IMI2	Innovative Medicines Initiative
IVD	in vitro diagnostic
LITMUS	Liver Investigation: Testing Marker Utility in Steatohepatitis
LOI	Letter of Intent
NAFLD	Non-Alcoholic Fatty Liver Disease
NAS	NAFLD Activity Score

NASH	Non-Alcoholic SteatoHepatitis
NASH-CRN	Non-Alcoholic SteatoHepatitis Clinical Research Network
QC	Quality Control
SAE	Serious Adverse event
SOP	Standard Operating Procedure
TMB	Tetramethylbenzidine
ROS	Reactive Oxygen Species
RUO	Research use Only
UK	United Kingdom
US	United States
VCTE	Vibration Controlled Transient Elastography

2. Administrative Information

Submission Title

To qualify a composite biomarker within the context of use (COU) prognostic enrichment

Requesting Organization

LITMUS (www.litmus-project.eu) is an EU funded consortium within the Innovative Medicines Initiative 2 (IMI2) Program (www.imi.europa.eu).
The Project Coordinator is Prof Quentin M. Anstee from Newcastle University, UK
Physical Address: Institute of Cellular Medicine, 4th Floor, William Leech Building, the Medical School, Framlington Place, Newcastle University, Newcastle-Upon-Tyne, NE2 4HH
Website: https://litmus-project.eu/

Specific information on the LITMUS consortium can be found in **Attachment 1.**

Biomarker information and COU

This LOI includes a composite biomarker consisting of 2 individual biomarkers. The two biomarkers are planned to be validated individually first before validating the final composite marker. The biomarkers are:

1: ELF™ can be measured in the serum as an indicator of liver fibrosis and cirrhosis

2: Iron corrected T1 (cT1) is a non-invasive biomarker, measured using magnetic resonance imaging (MRI)

3: The final composite biomarker, which will include both ELF™ and cT1, in this document defined as 'the final composite marker'

A list of questions is included to ensure that our current proposal is meeting the expectations of the FDA Biomarker Qualification Program (BQP). At this stage we would like to ask questions in relation to the use of the Metacohort for definition of cut-off values, the LITMUS study for validation of these cut-offs for the liquid biomarker (ELF™), and - for the Imaging biomarker (cT1) - using the first 100 patients of the Imaging study together with patients from 'the Oxford study' for definition of the cut-off values and the remaining approx. 400 patients of the Imaging study for validation of the cut-off. As the clinical studies are ongoing, it is of utmost importance for the LITMUS consortium to get feedback on these questions in case any changes would be required. We would appreciate the opportunity to implement such changes as soon as possible.

Be aware that the same clinical studies as well as the same consortium on December 5, 2019 submitted an LOI for the COU Diagnostic screening DDT No 000095.
Please note that some of the attachments are the same as were submitted to the FDA in the LOI for the COU Diagnostic screening DDT No 000095.

***References are available upon request*.**

Contact information

Physical Address of Regulatory Contact: Herlev Hovedgade 205-207, 2730 Herlev, Denmark
Phone Number: +45 2937 4027
Primary Point of contact: Elisabeth Erhardtsen (Direct: +45 2937 4027, email: eer@nordicbio.com), Nordic Bioscience A/S.

Back-up (keep in cc): Richard Torstenson, Allergan (Direct: +46 723220020, email: Richard.Torstenson@allergan.com); Quentin M. Anstee, Newcastle University (Direct: +44 (0) 191 208 7012, email: quentin.anstee@newcastle.ac.uk), Julia Brosnan, Pfizer (Direct: (860) 885-8394, Email: julia.brosnan@pfizer.com)

Submission Date

April 30 2020

3. Drug Development Need Statement

Describe the drug development need that the biomarker is intended to address, including (if applicable) the proposed benefit over currently used biomarkers for similar context of uses (COUs)

As recently acknowledged by the FDA[1]: *'NAFLD is the most common cause of chronic liver disease in North America. Currently, there are no approved drugs for the treatment of NASH. Given the high prevalence of NASH, the associated morbidity, the growing burden of end-stage liver disease, and limited availability of livers for organ transplantation, FDA believes that identifying therapies that will slow the progress of, halt, or reverse NASH and NAFLD will address an unmet medical need.'* and '*At this time, reliable diagnosis and staging of NASH can only be made by histopathological examination of a liver biopsy specimen. Liver biopsy, however, is an invasive procedure that is associated with occasional morbidity and, in rare circumstances, mortality. The use of liver biopsies in clinical trials poses significant logistical challenges (e.g., cost, availability of pathologists with specific expertise in NASH); in addition, some patients are reluctant or unwilling to undergo biopsy. Therefore, noninvasive biomarkers are needed (including imaging) to supplant liver biopsy and provide a comparable or superior ability to accurately diagnose and assess various grades of NASH and stages of liver fibrosis. Identification and validation of such biomarkers could significantly accelerate drug development in NAFLD. FDA encourages sponsors to consider biomarker development.'*
LITMUS is an EU funded consortium aiming to identify and validate biomarkers which can be qualified for use in the development of new therapies within NAFLD/NASH (see **Attachment 1**). Ultimately, the goal is to identify biomarkers that can decrease or eliminate the use of liver biopsy.

4. Biomarker Information and Interpretation

Biomarker names

4.1.1 The ADVIA Centaur® systems Enhanced Liver Fibrosis (ELF™) test

The ELF score is a unitless numerical value. The measuring units for the 3 constituent assays are in ng/mL. Limits of Detection (LoD) for the three assays are 1.6-1000 ng/mL for HA, 0.5-150 ng/mL for PIIINP, and 3.5-1300 ng/mL for TIMP-1. Please refer to **Attachment 2** for full details.

4.1.2 cT1

Liver iron corrected T1 (cT1) is a non-invasive, non-composite magnetic resonance biomarker that corresponds to the T1 relaxation time of liver tissue, correcting for the effects of hepatic iron content, as iron may result in an underestimation of liver disease by artificially shortening the T1. cT1 is measured in milliseconds, spanning a physiological range of 500-1400ms. Please refer to **Attachment 3** for full details**.**

Analytical methods

4.2.1 ELF™

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. The three individual constituent assays are fully automated, two-site sandwich assays using direct chemiluminometric technology.

4.2.2 cT1

T1 relaxation has been nicely summarised by Berger[2] who describes how magnetic resonance imaging (MRI) uses the body's natural magnetic properties to produce detailed images from any part of the body. For imaging purposes the hydrogen nucleus (a single proton) is used because of its abundance in water and fat. The hydrogen proton can be likened to behaving like a small bar magnet. Under normal circumstances, these hydrogen proton “bar magnets” spin in the body with their axes randomly aligned. When the body is placed in a strong magnetic field, such as an MRI scanner, the protons' axes all line up. This uniform alignment creates a magnetic vector oriented along the axis of the MRI scanner. When additional energy (in the form of a radio wave or radio frequency (RF) pulse) is added to the magnetic field, the magnetic vector is disturbed, and the protons are deflected from the uniform alignment. When the RF pulse is removed the nuclei return to their resting alignment and the time it takes for the protons to return to their resting equilibrium is called the longitudinal relaxation time (or T1). T1 is therefore a measure of how quickly protons re-equilibrate their spins after being excited by a radiofrequency pulse, which is dependent upon, and as a consequence a biomarker of, regional tissue water (proton) content. T1 is measured in milliseconds (ms) and is field strength dependent. Proton-dense tissues with a low water content, such as fat, have very short T1s, while tissues with a higher water content, such as muscle and the spleen, have much longer T1s. When tissue is inflamed or scarred (fibrotic), changes in the structural organization of the tissue, due to tissue remodeling, mean that the water content increases, leading to longer T1 values. cT1 refers to correcting the T1 signal for the presence of iron, a diamagnetic material, that opposes the MR signal. Iron content in the liver parenchyma can be measured using the MRI T2* relaxation time which is sensitive to local iron deposits. T2* and T1 are both measured in miliseconds. cT1 is also standardised across MR manufacturers and normalised to a single field strength. cT1 is measured in milliseconds, spanning a physiological range of 500-1400ms. Please refer to **Attachment 3**.

Measuring units and limits of detection

4.3.1 ELF™

The ELF score is a unitless numerical value. The measuring units for the 3 constituent assays are in ng/mL. Limits of Detection (LoD) for the three assays are 1.6-1000 ng/mL for HA, 0.5-150 ng/mL for PIIINP, and 3.5-1300 ng/mL for TIMP-1.

4.3.2 cT1

cT1 relates to the water content in liver parenchyma and is measured in milliseconds (ms). cT1 is standardised across MR manufacturers and normalised to a single field strength. cT1 measurement spans the physiological relevant range (500-1400ms).

4.3.3 The final composite marker

Not yet determined.

Biomarker Interpretation and utility

Post-analytical application/conversion of biomarker raw measure to the applied measure

4.4.1 ELF™

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). The ELF™ score can be calculated manually or by the ADVIA Centaur systems. The auto-calculation feature is only available for the ADVIA Centaur XPT system, for the ADVIA Centaur XP system software version 7.0 or higher, and for the ADVIA Centaur CP system software version 6.0 or higher. To calculate the ELF™ score manually for the ADVIA Centaur CP system, first obtain results for the HA, PIIINP, and TIMP-1 assays on the ADVIA Centaur CP system, and then use the following equation to calculate the ELF™ score: ELF™ score = 2.494 + 0.846 ln(HA) + 0.735 ln(PIIINP) + 0.391 ln(TIMP-1). Concentrations of each of the constituents are in ng/mL. The ELF™ score is a unitless numerical value. Please refer to **Attachment 2** for more details.

4.4.2 cT1

cT1 is a non-invasive, image derived biomarker used to indicate the free water within the tissue. cT1 data is acquired using the Liver*MultiScan* protocol and processed using the Liver*MultiScan* software. All MR data acquired with the Liver*MultiScan* protocol is uploaded to the custom-built secured data management system at Perspectum Diagnostics and processed by trained analysts. In brief, data undergoes automated quality control checks as well as visual inspection of parametric maps. On satisfactory maps, analysts either delineated the boundaries of the liver to derive an average value for whole liver segmentation maps using a semi-automatic method, or place three circular regions of interest (ROIs) with the default diameter of 15mm in representative areas of liver parenchyma on the chosen T2*, cT1 and PDFF maps. Please refer to **Attachment 3** for more details.

4.4.3 The final composite marker

Not yet determined.

Rationale for post-analytical elements

Describe rationale for post-analytical elements used as inputs in application or conversion of the raw biomarker measurement

4.5.1 ELF™

The ELF algorithm was identified in a 2004 study by Rosenberg et al[3]. The ELF score algorithm was initially determined and validated on the Immuno 1 analyzer (with associated reagents) and was refined over time and validated for use on the ADVIA Centaur systems.

4.5.2 cT1

The cT1 algorithm is the propriety property of Perspectum, and was developed in 2012 by Banerjee et al and further refined by Tunniclife et al (2017)[4]. cT1 was originally determined in a Siemens 3T scanner, but has since been refined to be acquired and standardised for acquisitions on Siemens, Phillips and GE scanners, and at two magnetic field strengths, 1.5 and 3 Tesla. See **Attachment 3**

4.5.3 The final composite marker

Not yet determined.

Clinical Interpretive Criteria

Describe the cut-off values, cut-points/thresholds, boundaries/limits or other comparators used in the interpretation of the biomarker measurement or its applied/converted form to draw an actionable conclusion based on the biomarker result

The cut-off values, cut-points/thresholds, or boundaries/limits for ELF™, cT1 and for the final composite marker are not yet determined. Optimal cut-offs for the biomarkers will for ELF™ be determined in the Metacohort Study (see **Attachment 4**).

For cT1 the optimal cut-off to identify those at the greatest risk of developing clinical relevant outcomes will be derived from a NAFLD cohort based on a combination of the data from the Jayaswal (Oxford-Reading) study (see Appendix 4 Training Data found in **Attachment 3**) and the first 100 patients included in the LITMUS Imaging study (see **Attachment 5**). In brief, both studies recruit(ed) patients suspected of NAFLD and referred for clinical biopsy.

Patients were/will be followed up for clinical events using patient medical records. For both biomarkers the approach is described in **Section 6** below, guided by statistical considerations reflected in the statistical analysis plan (see **Attachment 6**), thereafter validated for ELF™ in the LITMUS Study (see **Attachment 7**), and for cT1 the validation will be performed in the remaining patients in the Imaging Study (see **Attachment 5**)**.**

5. Context of Use Statement (COU)

The proposed prognostic enrichment biomarkers (liver cT1 and/or the ELF test) should identify patients who are more likely to reach the intermediate and/or the composite long-term events. For this COU the long term events are defined as Death (liver-related or all-cause), liver transplant, complications of cirrhosis (including hepatocellular carcinoma (HCC); variceal bleed; Change in MELD score from less than or equal to 12 to more than 15, and histological progression to cirrhosis. The intermediate endpoint will be defined based on the definition in the reflection paper below.

Such prognostic biomarkers can increase the likelihood of patients reaching the endpoints, and thereby reduce the number of patients to be included in trials. Consequently, the final trial will have an increased number of clinical events and thereby target the group of NASH patients with a high medical need. This should also expedite the development of new therapies in active NASH, as fewer patients will be needed when a higher event rate is present.

Additionally, the exclusion of patients that reach the intermediate endpoints without treatment from clinical trials, i.e. those that spontaneously regress in disease, will also increase the efficiency of such trials. FDA has accepted as critical inclusion criteria in NASH trials a NASH activity score greater than or equal to 4 with at least 1 point each in inflammation and ballooning *along with* a NASH Clinical Research Network (CRN) fibrosis score greater than stage 1 fibrosis but less than stage 4 fibrosis. These two criteria ensure that patients have evidence of steatohepatitis and significant liver fibrosis without cirrhosis at enrolment. Depending on the drug's mechanism of action and anticipated effect on inflammation and/or fibrosis, the sponsor can propose for discussion with the FDA alternatives to the NAS and NASH/CRN fibrosis score. The sponsor should provide adequate scientific justification for the alternatives[1].

The paper also defined the intermediate endpoints as well as a composite long-term endpoints

'The intermediate endpoints are defined as:

Resolution of steatohepatitis on overall histopathological reading **and** no worsening of liver fibrosis on NASH CRN fibrosis score. Resolution of steatohepatitis is defined as absent fatty liver disease or isolated or simple steatosis without steatohepatitis and a NAS score of 0–1 for inflammation, 0 for ballooning, and any value for steatosis; OR

— Improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis 300 score) **and** no worsening of steatohepatitis (defined as no increase in NAS for 301 ballooning, inflammation, or steatosis); OR

— Both resolution of steatohepatitis and improvement in fibrosis (as defined above).

The composite long-term endpoint is defined based on the following events (whichever comes first):

Progression to cirrhosis on histopathology. Reduction in hepatic decompensation events (e.g., hepatic encephalopathy, variceal bleeding, ascites). These events should be adjudicated by a committee of experts. Change in MELD score from less than or equal to 12 to more than 15. (This endpoint approximates listing for liver transplant). Liver transplant. All-cause mortality.

- **General Area**: A prognostic enrichment biomarker that would create a clinical trial population more likely to progress in disease. The preferential inclusion of patients more likely to progress would make future trials in NASH patients more efficient, requiring fewer participants to obtain a similar power, or increasing power with the same number of participants.
- **Target Population for use**: NAFLD/NASH patients
- **Stage of Drug Development for Use**: All clinical stages of NASH drug development, including proof of concept, dose-ranging, and confirmatory clinical trials.

A decision tree for the prognostic enrichment biomarkers is presented in **Attachment 8.**

6. Analytical Considerations

General description of what aspect of the biomarker is being measured including Index scoring as appropriate

ELF™

Longitudinal serum specimens will be used for the ELF™ test. The full panel of molecular targets was selected to include 9 surrogate markers of matrix synthesis or degradation, based on knowledge of the basic mechanisms involved in liver fibrosis. From this panel, the combination of HA, PIIINP and TIMP-1 was shown to perform best in the detection of liver fibrosis[3]. For the ADVIA Centaur XP and XPT system, the ELF™ score is calculated using the following equation: ELF™ score = 2.278 + 0.851 ln(CHA) + 0.751 ln(CPIIINP) + 0.394 ln(CTIMP-1). Please refer to **Attachment 2. section 3.4.b**

cT1

cT1 data is acquired using the Liver*MultiScan* protocol and processed using the Liver*MultiScan* software. All MR data acquired with the Liver*MultiScan* protocol is uploaded to the custom-built

secured data management system at Perspectum and processed by trained analysts. In brief, data undergoes automated quality control checks as well as visual inspection of parametric maps. On satisfactory maps, analysts either place three circular regions of interest (ROIs) with the default diameter of 15mm in representative areas of liver parenchyma on the cT1 maps or delineate the boundary of the liver using semi-automated liver segmentation tool. Pixel values in the ROIs or in the entire segmentation mask are then averaged and the median value used as a representative measure for the overall liver tissue status.

Description of sample source, matrix (base material and any additives), stability and composition of biomarker

The sample source for ELF™ will be selected from the Metacohort while for cT1, data acquired as part of the Oxford-Reading study (see Appendix 4 Training Data found in **Attachment 3**) and the first 100 patients included in the LITMUS Imaging study (see **Attachment 5**) will be used as a training set (see Appendix 4 Training Data found in **Attachment 3**). For these biomarkers it is not the intention to select cut-off points but rather evaluate the prognostic performance of both markers.
Samples for validation of the prognostic performance of ELF™ will be based on samples from the LITMUS study (see **Attachment 7**). For the validation of cT1 prognostic performance measurements will be obtained according to the imaging protocol (see **Attachment 5).**
The biospecimens from the Metacohort have been stored at -80°C for up to 18 years; most have been thawed and refrozen one or two times. The prospectively collected samples from the LITMUS study have been stored at -80°C according to the sampling protocol (see **Attachment 9**).

6.3.1 ELF™ sample stability and robustness

The ADVIA Centaur ELF test provides 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). The ADVIA Centaur ELF test has been validated for use with human serum specimens. The stabilities of the constituent analytes of the ADVIA Centaur ELF test have been evaluated under various storage conditions. The current recommended storage conditions are presented below:

1. Do not use samples that have been stored at room temperature for longer than 48 hours.
2. Tightly cap and refrigerate samples at 2–8°C if the assay is not completed within 48 hours.
3. Specimens may be stored on the clot.
4. Freeze samples at or below -20°C, if the sample is not assayed within 7 days.
5. Do not store in frost-free freezer.
6. Freeze samples devoid of red blood cells up to 4 times, and mix thoroughly after thawing.
7. Centrifuge thawed samples at 1000 x g for 10 minutes before using.

It has been reported in literature that samples are stable following freeze-thaw cycles, refrigeration over days at 4°C, and long-term storage at -80°C. In a study by Kennedy et al[5], recoveries of the individual analytes were observed to change by up to ±20%. However, in these samples, the mean ELF™ score did not vary by >0.1 units (0.7%) and the maximum observed change in ELF™ was 0.2 units[5]. Separate studies by Puigvehí et al[6] and Dellavance et al[7] showed that the ELF™ score remained stable in samples stored over a period of 20 years and samples subjected to 9 freeze/thaw cycles, respectively[6,7].

6.3.2 cT1

There are no sample degrading issues associated with cT1 as it is derived from MRI data stored digitally in DICOM format. Furthermore, the results of the histological reads and clinical events are stored electronically in an eCRF.

6.3.3 The final composite marker

Will consist of ELF™ and cT1, see above.

6.4 Description of factors and plans to preserve specimen integrity

6.4.1Description of pre-analytical factors and quality assurance/quality control (QA/QC) plans to preserve specimen integrity

A standard operating procedure (SOP) for sample collection including timing and location that sample will be collected from, storage and test/assay methodology; reference or control samples.

For the LITMUS Study the quality of samples and measurements is ensured by collecting all data according to the protocol and sampling protocols (see **Attachment 7** and **9**). Furthermore, monitoring plans for both the LITMUS Study and the Imaging Study are in place (see **Attachment 10** and **11**), as well as a Data management plan (see **Attachment 12**). Moreover, measuring of ELF samples is done in a CLIA certified laboratory.
For the Metacohort Study, no specific monitoring plan has been in place, outside good laboratory practice (see **Attachment 4**). ELF levels from samples of the Metacohort study are analyzed in a CLIA certified laboratory. The ELF test will be performed according to manufactures IFU, the analysis is at present for RUO.

6.4.2 Analytical validation plan

Analytical validation plan: description of measurement tool and device calibrations

6.4.2.1 Technical performance validation ELF™

The ADVIA Centaur ELF™ test has been validated according to the protocols described in the White Paper[8]. The ADVIA Centaur test is CE marked and received Breakthrough designation in the US. The ELF™ test is for Research Use Only.

6.4.2.2 Technical performance cT1

Technical validity of cT1 has been demonstrated in terms of accuracy (using MRI Phantom data) and repeatability and reproducibility of the signal (using Phantom and Human data). In vivo performance testing to assess the repeatability and reproducibility of cT1 using volunteers with a range of physiological values has been conducted. Repeatability in terms of the variability across device measurements under the same measurement conditions (volunteer on/off/on), and reproducibility as the variation across device measurements between reference and non-reference scanners. Full results have been published (Bachtiar, 2019[9]) and can be found in Appendix 2 Measuring Iron Corrected T1 - RA252 found in **Attachment 3**. cT1 has also been demonstrated to have operator independence (See Appendix 3 operator reliability found in **Attachment 3**)

6.4.2.3 The final composite score

See above - as this will consist of ELF™ and cT1

6.5 Validation of the final version of the measurements tool

Once the SOP and analytical validation plan is finalized, describe how you will use this process to validate the final version of the measurement tool

The results of the technical validation of the individual biomarkers as well as the clinical performance (cut-off defined in Metacohort study and the cT1 training data (as described above) and confirmed in the LITMUS study and, for cT1, in the Imaging study) will be the basis for the final composite biomarker for qualification.

7. Clinical considerations

Describe how the biomarker measurement is used to inform drug development. Please provide a decision tree to guide how the biomarker information would be used in drug development or a clinical trial.

COU Prognostic – Enrichment

The decision tree for prognostic enrichment is shown in **Attachment 8**. Individuals are screened for the initial eligibility criteria. If the patient meets the eligibility criteria, the patient will have the final composite marker measured. If the values are above the selected cut-off, patients are at high risk of developing events (defined in section 5), and the individual will be enrolled in the clinical trial. A prognostic enrichment biomarker can enrich a trial by including individuals more likely to develop an event such as progression to fibrosis/developing advanced fibrosis. This will increase the power and potentially shorten clinical trials.

Describe patient population or drug development setting in which the biomarker will be used.

Non-alcoholic fatty liver disease (NAFLD) is a common progressive disorder closely associated with the clinical features of metabolic syndrome. This chronic liver condition occurs by excessive accumulation of fatty acids within hepatocytes and also represents a range of alterations to the extracellular matrix[10–12]. NAFLD represents a wide spectrum of disease ranging from simple steatosis to NASH, which is characterized by hepatic steatosis, inflammation and hepatocyte injury with variable degrees of fibrosis in the absence of secondary causes of steatosis[11,13]. Changes in the architecture and composition of the liver have been shown to be associated with clinically relevant progression of the disease[14–16]. The biomarkers proposed herein reflect various aspects of the alterations taking place in the liver during disease progression.

Clinical validation: provides information to support biological and clinical relevance of the biomarker as applied in the COU.

The clinical validation for each of the two biomarkers will be done utilizing two separate NASH/NAFLD study groups.

The Metacohort Study (see **Attachment 4**) will be used to define cut-offs for ELF™.

For cT1 the optimal cut-off to identify those at the greatest risk of developing clinical relevant outcomes will be derived from a NAFLD cohort based on a combination of the data from the Jayaswal (Oxford-Reading) study (see Appendix 4 Training Data found in **Attachment 3**) and the first 100 patients included in the LITMUS Imaging study (**see Attachment 5**).

The LITMUS study is performed under a protocol, patients will be included according to inclusion

and exclusion criteria as described in the LITMUS Study protocol (see **Attachment 7**), using the technically validated biomarker **ELF™** which will be validated according to the white paper released by the Biomarker Assay Collaborative Evidentiary Considerations Writing Group, Critical Path Institute (C-Path)[8].

cT1 will be validated as described in the imaging study protocol (a substudy of the LITMUS study;see **Attachment 5**).

11 countries in Europe will be involved in the LITMUS Study.

7.3 Describe how normal or other reference values are established, provide study design(s), analytical plan, etc.

7.3.1 Reference value

Two reference endpoints will be described. The first reference endpoint is related to the histological assessment of liver biopsy, which is part of both the intermediate endpoint, as well as the long-term endpoints. The second reference endpoint relates to the long-term outcome of the patients and the development of events such as death (liver-related or all-cause), liver transplant, complications of cirrhosis (including HCC, variceal bleeding), change in MELD score from less than or equal to 12 to more than 15.

These events will be collected for ELF in the Metacohorte as well as the LITMUS study, while for cT1 the events will be collected in the Oxford study as well as the Imaging study (see **Attachment 4**, **5**, **7**). The events will be collected based on either SAE forms, alternatively for studies in which these are collected as events and not as SAE they will be captured via the Patient record forms.
The LITMUS Study and the Imaging sub study are prospective clinical studies for which data are collected in accordance with clinical protocols (see **Attachment 5** and **7**), a sampling protocol (see **Attachment 9**), quality is ensured by monitoring plans (see **Attachment 10** and **11**), a standardized procedure for obtaining, processing and assembly of a liver biopsy in the LITMUS study (see **Attachment 13**), and a data management plan (see **Attachment 12**). Furthermore, the imaging charter can be found in **Attachment 14**. Statistical analysis to be performed is described in the statistical analysis plan (see **Attachment 6**).

7.3.2 Benefit and Risk

Benefits and Risks of applying the biomarker in drug development or a clinical trial.
A liver biopsy has an inherent risk of discomfort, bleeding, and in very rare cases death. A qualified biomarker for the described COU prognostic enrichment should decrease the number of biopsies performed by enriching the trial population. At this point in the development the risks have not been identified.

7.3.3 Knowledge gaps

Describe any current knowledge gaps, limitations and assumptions in applying the biomarker in drug development or a clinical trial
At this point the knowledge gaps, limitations and assumptions for utilization of the mentioned biomarker have not been defined.

8. Supporting Information

8.1.1 Provide underlying biological process or supporting evidence of association of the biological process with the biomarker

Liver fibrosis is biochemically complex but is orchestrated primarily by activated hepatic stellate cells (HSCs). Activated HSCs produce components of the extracellular matrix (ECM). The ECM includes an array of proteins involved in scar formation including fibronectin, laminin, collagens, hyaluronic acid (HA), and proteoglycans. Type I, III, IV, and V collagen are prominently expressed within the liver[17]. HA is an essential component of the ECM and is produced primarily by HSC[18]. The accumulation of deposited ECM progressively replaces the normal liver parenchyma, producing damage and scar tissue and ultimately disrupting hepatic architecture and function.

8.1.2 ELF™ biology

Fibrosis of the liver is a largely bidirectional process[19,20]. Both fibrosis and repair mechanisms have been linked to ECM-related pathways. Regression and repair are associated with upregulation of matrix metalloproteinases (MMPs), which are a family of zinc-dependent endopeptidases capable of degrading ECM deposition and so central 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[20,21]. Additionally, low levels of TIMP-1 may promote hepatic stellate cell apoptosis[19]. 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. N-terminal pro-peptide of procollagen type III (PIIINP) is generated during the synthesis of type III collagen. PIIINP can be measured in the serum as an indicator of liver fibrosis and cirrhosis[3,22].

The components of the ELF test were identified in a 2004 study by Rosenberg et al. The full panel of molecular targets was selected to include 9 surrogate markers of matrix synthesis or degradation, based on knowledge of the basic mechanisms involved in liver fibrosis. From this panel, the combination of HA, PIIINP and TIMP-1 were shown to be the most useful in the detection of liver fibrosis[3].

For the ADVIA Centaur XP and XPT system, the ELF score is calculated using the following equation:

$$\text{ELF score} = 2.278 + 0.851 \ln(\text{CHA}) + 0.751 \ln(\text{CPIIINP}) + 0.394 \ln(\text{CTIMP-1})$$

For the ADVIA Centaur CP system, the ELF score is calculated using the following equation:

$$\text{ELF score} = 2.494 + 0.846 \ln(\text{CHA}) + 0.735 \ln(\text{CPIIINP}) + 0.391 \ln(\text{CTIMP-1})$$

More details on the rationale of parameter selection are provided in **Attachment 2**.

8.1.3 cT1

T1 relaxation has been nicely summarised by Berger[2] who describes how magnetic resonance imaging (MRI) uses the body's natural magnetic properties to produce detailed images from any part of the body. For imaging purposes the hydrogen nucleus (a single proton) is used because of its abundance in water and fat. The hydrogen proton can be likened to behaving like a small bar magnet. Under normal circumstances, these hydrogen proton “bar magnets” spin in the body with their axes randomly aligned. When the body is placed in a strong magnetic field, such as an MRI scanner, the protons' axes all line up. This uniform alignment creates a magnetic vector oriented along the axis of

the MRI scanner. When additional energy (in the form of a radio wave or radio frequency (RF) pulse) is added to the magnetic field, the magnetic vector is disturbed and the protons are deflected from the uniform alignment. When the RF pulse is removed the nuclei return to their resting alignment and the time it takes for the protons to return to their resting equilibrium is called the longitudinal relaxation time (or T1). T1 is therefore a measure of how quickly protons re-equilibrate their spins after being excited by a radiofrequency pulse and as such is an indicator of regional tissue water (proton) content. T1 is measured in milliseconds (ms) and is field strength dependent. Proton-dense tissues with a low water content, such as fat, have very short T1s, while tissues with a higher water content, such as muscle and the spleen, have much longer T1s. When tissue is inflamed or scarred (fibrotic), changes in the structural organization of the tissue, due to tissue remodelling, mean that the water content increases, leading to longer T1 values.

More details on the rationale of parameter selection are provided in **Attachment 3**

Summary of existing clinical data to support the biomarker in its COU (e.g. summaries of literature findings, previously conducted studies).

8.2.1 ELF – Clinical data as a prognostic enrichment marker

See **Attachment 2**

8.2.2 cT1 – Clinical data as a prognostic enrichment marker

See **Attachment 3**

8.2.3 – Clinical data on the final composite score

Clinical data on the final score has not yet been obtained.

9. Previous Qualification Interactions and Other

Qualification Interactions

An Innovation Task Force briefing meeting was held at the European Medicines Agency (EMA) on 10th October 2018. This was a general meeting at which the individual biomarkers were not discussed.

On June 15th 2019 a scientific advice was started on qualification of markers for Diagnostic Screening (PRO-C3 and the FAST score (Fibroscan and AST)) and another scientific advice on Prognostic Enrichment (ELF and cT1). A meeting was held on February 11th 2020 with the defined group at EMA premises, written feedback was received on April 6th 2020.

On June 16th 2019 an LOI was submitted to FDA, which included 3 different COU in one LOI, and on July 17th 2019 a phone meeting was held with the FDA qualification team. At this meeting LITMUS was requested to send one COU per LOI, and informed that each LOI can only include one biomarker - however a biomarker which is a composite was mentioned as acceptable. An updated LOI on Diagnostic screening was submitted on December 4 2020 DDT No 000095, this has been accepted for review on February 27 2020.

The current LOI on Prognostic Enrichment was written in response to the request on July 17th 2019.

Other regulatory interactions ELF™

Interactions with the FDA and EMA on ELF™

Discussions with FDA regarding Breakthrough Device Designation and relevant follow-up discussions were completed under Q181316 and applicable supplements. Using self-declaration in order to be placed on the market. No discussions initiated by Siemens have been held with EMA.

IFU:

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. The ADVIA Centaur ELF test is indicated, in conjunction with other laboratory findings and clinical assessments, as an aid in the diagnosis and assessment of the severity of liver fibrosis in patients with signs and symptoms of chronic liver disease. This test is not intended for use on any other system.

FDA	Activity	Reviewing Office	Date	Status
ELF test	Breakthrough Device Designation	CDRH	Aug. 20, 2018	Q181316 and applicable supplements
ELF test	De Novo Requests	CDRH	Jan. 2, 2020	In review
ELF test	DDT-Qualification process Through NIMBLE consortium	CDER	x	LOI accepted

EMA	Activity	Notified bodies	Competent Authority	Date	Status
ELF test	CE marked	NA	MHRA/HPRA	2011	CE Marked

Other regulatory interactions on cT1

Interactions with the FDA and EMA on cT1

FDA	Activity	Reviewing Office	Date	Status
Liver*MultiScan* device (in which cT1 is described as a biomarker)	510k pre-market notification	CDRH	Nov 2017	Cleared
cT1 biomarker	Diagnostic screening briefing doc Diagnostic screening transition Validation study review $250k grant for ongoing BQP activities	CDER CDER CDER CDER	Jan 2018 Jan 2019 Aug 2019 Sept 2019	Submitted Support for QP Study approval Awarded

EMA	Activity	Notified bodies	Competent Authority	Date	Status
Liver*MultiScan* device	CE Mark	BSI	MHRA	Feb 2018	Approved
cT1 biomarker	Diagnostic screening in NASH	N/A	EMA	Mar 2019 Sept 2019	Submitted Letter of support

10. Attachments

Attachment 1 = LITMUS charter
Attachment 2 = ELF information
Attachment 3 = cT1 (contains 4 sub-appendices)
Attachment 4 = Metacohort Study description
Attachment 5 = Imaging study protocol (LITMUS substudy)
Attachment 6 = Statistical analysis plan LITMUS
Attachment 7 = LITMUS Study protocol
Attachment 8 = Decision tree for Prognostic enrichment
Attachment 9 = Sampling protocol (Study Handbook)
Attachment 10 = Monitoring plan for the LITMUS study
Attachment 11 = Monitoring plan for the Imaging study
Attachment 12 = Data management plan
Attachment 13 = Biopsy description
Attachment 14 = Imaging charter

11. Questions for BQP

Questions on clinical validation of biomarkers:

- Is the Metacohort Study (see **Attachment 4)** sufficient to be used for determination of cut-offs for ELF™ (Note: has been technically validated up to CE level before measuring the individual levels in this patient cohort)?

- The plan is to define the cut-offs based on data from the first 100 patients enrolled in the Imaging study along with the patients from the Oxford study. Is this number considered sufficient for definition of cut-offs for the COU diagnostic screening.
- Do the protocols for the prospective LITMUS Study (see **Attachment 7**) and the Imaging Study (see **Attachment 5**), from which data will be collected for validation of the cut-offs, correctly define the primary objective, inclusion/exclusion criteria, SAE/AE reporting and sample collection, for a later qualification of the final composite biomarker for the COU diagnostic screening? (The ELF™ test to be used at this step has been validated according to the White paper[8]).
- **Plans, SOPs, SAP and monitoring of the prospective LITMUS study**:
    - Reference markers. All biomarkers will be compared to liver biopsies and the Liver related events, is the biopsy description (see **Attachment 13**) and the composite endpoint see page correctly defined for a later qualification of the final composite biomarker?
    - Are the protocol and monitoring plans for the LITMUS Study (**Attachment 7** and **10)** and Imaging Study (**Attachment 5** and **11**) in accordance with FDA expectations for data that will be used for a later qualification of the final composite biomarker?
    - Is the data management plan (see **Attachment 12**) considered sufficient for a later qualification of the final composite biomarker?
    - Is the SOP for collection, processing, and storage of plasma, serum and frozen biopsies considered sufficient (**Attachment 9**) for a later qualification of the final composite biomarker?

12. References

1. FDA/CDER. Noncirrhotic Nonalcoholic Steatohepatitis With Liver Fibrosis: Developing Drugs for Treatment Guidance for Industry DRAFT GUIDANCE. (2018).
2. Berger, A. Magnetic resonance imaging. *BMJ* **324,** 2002 (2002).
3. Rosenberg, W. M. C. *et al.* Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* **127,** 1704–13 (2004).
4. Tunnicliffe, E. M., Banerjee, R., Pavlides, M., Neubauer, S. & Robson, M. D. A model for hepatic fibrosis: the competing effects of cell loss and iron on shortened modified Look-Locker inversion recovery T1 (shMOLLI-T1) in the liver. *J. Magn. Reson. Imaging* **45,** 450–462 (2017).
5. Kennedy, O. J., Parkes, J., Tanwar, S., Trembling, P. M. & Rosenberg, W. M. The Enhanced Liver Fibrosis (ELF) Panel: Analyte Stability Under Common Sample Storage Conditions Used in Clinical Practice. *J. Appl. Lab. Med. An AACC Publ.* **1,** 720–728 (2017).
6. Puigvehí, M. *et al.* Diagnostic Accuracy of the Enhanced Liver Fibrosis (ELF®) Score Using HCV-Infected Serum Samples Cryopreserved for up to 25 Years. *PLoS One* **11,** e0164883 (2016).
7. Dellavance, A. *et al.* Enhanced liver fibrosis (ELF) score: Analytical performance and distribution range in a large cohort of blood donors. *Clin. Chim. Acta.* **461,** 151–5 (2016).
8. Piccoli, S. *et al.* Points to Consider Document: Scientific and Regulatory Considerations for the Analytical Validation of Assays Used in the Qualification of Biomarkers in Biological Matrices. *Biomark. Assay Collab. Evidentiary Considerations Writ. Group, Crit. Path Inst.* 1–54 (2017).
9. Bachtiar, V. *et al.* Repeatability and reproducibility of multiparametric magnetic resonance imaging of the liver. *PlosOne* **10,** 1–12 (2019).
10. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic

fatty liver disease. *Hepatology* **41,** 1313–1321 (2005).
11. EASL-EASD-EASO. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J. Hepatol.* **64,** 1388–1402 (2016).
12. Schuppan, D., Surabattula, R. & Wang, X. Y. Determinants of fibrosis progression and regression in NASH. *J. Hepatol.* **68,** 238–250 (2018).
13. Chalasani, N. *et al.* The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* **67,** 328–357 (2018).
14. Hagström, H. *et al.* Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J. Hepatol.* **67,** 1265–1273 (2017).
15. Dulai, P. S. *et al.* Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. *Hepatology* **65,** 1557–1565 (2017).
16. Ekstedt, M. *et al.* Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* **61,** 1547–1554 (2015).
17. Martinez-Hernandez, A. The hepatic extracellular matrix. I. Electron immunohistochemical studies in normal rat liver. *Lab. Invest.* **51,** 57–74 (1984).
18. Orasan, O. H., Ciulei, G., Cozma, A., Sava, M. & Dumitrascu, D. L. Hyaluronic acid as a biomarker of fibrosis in chronic liver diseases of different etiologies. *Med. Pharm. Reports* **89,** 24–31 (2016).
19. Ramachandran, P. & Iredale, J. P. Liver fibrosis: a bidirectional model of fibrogenesis and resolution. *QJM* **105,** 813–817 (2012).
20. Yoshiji, H. *et al.* Tissue inhibitor of metalloproteinases-1 attenuates spontaneous liver fibrosis resolution in the transgenic mouse. *Hepatology* **36,** 850–860 (2002).
21. Yilmaz, Y. & Eren, F. Serum biomarkers of fibrosis and extracellular matrix remodeling in patients with nonalcoholic fatty liver disease. *Eur. J. Gastroenterol. Hepatol.* **31,** 43–46 (2019).
22. Tanwar, S. *et al.* Validation of terminal peptide of procollagen III for the detection and assessment of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease. *Hepatology* **57,** 103–111 (2013).