DERIVATION AND PERFORMANCE OF STANDARDISED ENHANCED LIVER FIBROSIS (ELF) TEST THRESHOLDS FOR THE DETECTION AND PROGNOSIS OF LIVER FIBROSIS.

James Day¹, Preya Patel¹, Julie Parkes¹,² and William Rosenberg¹

¹The Institute for Liver and digestive Health, UCL Division of Medicine, UCL, London, UK
²The Department of Public Health Sciences and Medical Statistics, University of Southampton, UK

Corresponding Author:
Prof William Rosenberg,
Institute for Liver and Digestive Health,
UCL Medical School,
Royal Free Hospital
Pond Street
London NW3 2PF

Telephone:+44 (0) 7711016708
Fax: +44 (0) 2074726226
e-mail: w.rosenberg@ucl.ac.uk
Conflict of interest: JD and PP none. WMR is an inventor of the ELF test. He has received fees from Siemens Healthineers for lecturing, travel and consulting and research support. JP has received fees from Siemens Healthineers for lecturing.

Financial support: The original EUROGOLF study was supported by Bayer Pharma and the follow-up for clinical outcomes was supported by the MRC and Siemens Healthineers.
Authors’ contributions: WMR and JP conceived the study, designed it and contributed to the analysis of the data and writing. JD and PP contributed to data analysis and writing the manuscript.
Abstract

**Introduction:** Non-invasive tests are increasingly used to assess liver fibrosis and determine prognosis but suggested test thresholds vary. We describe the selection of standardised thresholds for the Enhanced Liver Fibrosis (ELF) test for the detection of liver fibrosis and for prognostication in chronic liver disease.

**Methods:** A Delphi method was used to identify thresholds for the ELF test to predict histological liver fibrosis stages including cirrhosis, using data derived from 921 patients in the EUROGOLF cohort. These thresholds were then used to determine the prognostic performance of ELF in a subset of 457 patients followed for a mean of 5 years.

**Results:** The Delphi panel selected sensitivity of 85% for the detection of fibrosis and >95% specificity for cirrhosis. The corresponding thresholds were 7.7, 9.8 and 11.3. Eighty five percent of patients with mild or worse fibrosis had an ELF score ≥7.7. The sensitivity for cirrhosis of ELF≥9.8 was 76%. ELF ≥11.3 was 97% specific for cirrhosis. ELF scores show a near linear relationship with Ishak fibrosis stages. Relative to the <7.7 group, the hazard ratios for a liver related outcome at 5 years were 21.00 (95% CI 2.68-164.65) and 71.04 (95% CI 9.4-536.7) in the 9.8 to <11.3 and ≥11.3 sub-groups respectively.

**Conclusion:** The selection of standard thresholds for detection and prognosis of liver fibrosis is described and their performance reported. These thresholds should prove useful both in interpreting and explaining test results and when considering the relationship of ELF score to Ishak stage in the context of monitoring.
Lay Summary

Using data derived from a large prospective study and the opinions of expert hepatologists we have identified standard thresholds for the Enhanced Liver Fibrosis test that can be used to detect liver fibrosis of different degrees of severity, and determine the prognosis of chronic liver disease.
**Introduction**

In most chronic liver diseases liver fibrosis is a consequence of liver injury with the accumulation of collagenous scar tissue eventually leading to cirrhosis with distortion of the hepatic vasculature. Cirrhosis, the fifth commonest cause of death in the UK in middle aged men, is still on the increase and is thus a major health problem(1). The prognosis of liver disease varies dramatically as cirrhosis advances, with the 1 year mortality rate in decompensated disease as high as 57%(2). The morbidity and mortality associated with cirrhosis can be reduced with appropriate interventions. It is therefore crucial to detect cirrhosis as early as possible in order to treat patients in the hope of reducing the incidence of the complications of decompensated disease including oesophageal varices, hepatic encephalopathy and ascites(3, 4) and to manage early hepatocellular cancers(5).

Histological analysis of liver biopsies has been the established reference standard for the assessment of liver fibrosis(6), and is usually considered the most specific test to determine the severity of liver disease(7). Clinicians readily accept the staging of a liver biopsy as indicative of mild, moderate or severe fibrosis and use numerical staging systems that assign numbers (0-6 or 0-4) to describe these stages of fibrosis. Increasing recognition of the hazards associated with biopsy and limitations in the diagnostic performance of histological staging have led to the quest for non-invasive tests for liver
fibrosis that might augment or replace the use of liver biopsies to stage disease severity and determine prognosis (8-11).

Non-invasive tests for liver fibrosis include blood tests, fibro-elastography and imaging. These tests can be standardized to yield good analytical and diagnostic performance, with reliable reproducibility, and may be repeated frequently. The information generated in the form of a numerical fibrosis score can be related to both the severity of liver fibrosis and the prognosis for clinically significant outcomes. These continuous variable scores reflect the biological process of liver fibrosis more accurately than categorical histological stages.

Both elastography and imaging have been shown to perform well in the detection of severe fibrosis and cirrhosis (12, 13) but less well in mild disease (14). Both are dependent on access to well-maintained equipment, skilled operators and interpreters.

Blood tests have the advantages that samples can be obtained easily, safely, and relatively painlessly. Analytical methods range from the calculation of scores based on the measurement of routine biochemical and haematological tests such as APRI (15), FIB4 and the Forn’s index (16-19); to panels including less routine biochemical tests such as FibroTest (20); to those that incorporate the measurement of constituents of fibrous matrix such as Hepascore (21) and Fibrometer (22). These tests tend to perform better at differentiating between severe and mild fibrosis rather than precisely distinguishing between each
histological stage(23, 24), but can provide useful prognostic information(25). However, many of the studies focus on the performance of the test in one particular disease rather than taking a range of aetiologies and most are affected by liver inflammation. Combinations of biomarker panels with imaging may improve the diagnostic accuracy in distinguishing between stages of fibrosis(26-28) but carry the disadvantages associated with physical tests.

The Enhanced Liver Fibrosis (ELF) test is a serum test that measures hyaluronic acid (HA), procollagen III amino acid terminal peptide (PIIINP) and tissue inhibitor of metalloproteinase 1 (TIMP-1); three molecules directly involved in liver matrix metabolism. The ELF test was derived and validated in separate cohorts selected at random from over 1,000 patients recruited in the EUROGOLF study reported in 2004(29). Subsequently the ELF test has been validated in a wide range of chronic liver diseases for the assessment of liver fibrosis(29, 30), including the paediatric population(31), and has been shown to be an accurate prognostic marker for all-cause mortality and the complications of cirrhosis(32) and may be more accurate than liver histology in determining prognosis(33).

The ELF test is marketed by Siemens Healthineers. The instructions for use provided by the manufacturer with the ELF reagents recommend test thresholds that can be used to interpret the results. However the data underpinning these recommended have not previously been placed in the public domain. Here we describe the process that was undertaken to identify these thresholds for the ELF test and their relationship to histologically
defined stages of fibrosis through analysis of test performance in samples from the cohorts of patients in which the test was derived and validated. These data were shared with the manufacturer and used to establish the thresholds described in their instructions for use.

In addition to the manufacturer’s thresholds we describe a new threshold for the detection of cirrhosis with high specificity.

Furthermore, in order to exploit the additional information captured by a continuous variable such as ELF, we have investigated the change in ELF score within each histological stage to describe the relationship between the ELF score and liver fibrosis staging.

Although the prognostic performance of ELF has been described in specific and diverse chronic liver diseases(32-34), the thresholds identified in these studies were derived from study specific data rather than the manufacturer’s thresholds. In order for clinicians to interpret both fibrosis severity and prognosis using the same thresholds we have determined the prognostic performance of ELF in predicting liver-related morbidity and liver-related deaths in a mixed cohort of patients with chronic liver disease over 2, 5 and 6 years follow-up at the manufacturer’s thresholds, and our new threshold specific for cirrhosis.

**Materials and Methods**

Clinically significant sensitivities and specificities for the detection of liver fibrosis were determined through a Delphi process(35) in which four expert
hepatologists were interviewed by WR or JP and asked to select levels of
diagnostic performance that they would deem acceptable when assessing
patients with chronic liver disease. Specifically they were asked what
proportion of patients with mild or worse liver fibrosis they would be willing to
miss-assign as having “no fibrosis”; and the proportion of patients with severe
fibrosis or cirrhosis that they would be willing to miss-assign as having
“moderate or milder fibrosis” to the nearest 5%. After identifying these
sensitivities and specificities, the data generated in the original EUROGOLF
cohort(29) were interrogated to identify the ELF test thresholds. After
reviewing the performance of these thresholds, the hepatologists then
requested a further threshold that was even more specific for liver cirrhosis.

We have previously reported the event-free survival, and probability of liver
related morbidity and mortality, for the 501 English patients recruited in the
EUROGOLF cohort(33). These data were used a priori to determine
thresholds that correlated with categorical risks for liver related outcomes. The
thresholds generated using this approach differed from those derived from the
cross-sectional analysis of histological fibrosis severity as described above. In
order to avoid confusion arising from the use of different thresholds for
assessment and prognosis, we recalculated the prognostic performance of
the ELF test at the thresholds identified for histological assessment, using the
data obtained for the English participants in the EUROGOLF cohort.

The ELF prognostic performance study is based on follow up of patients
recruited to the original Bayer “EUROGOLF” study in 1998-2000 conducted
by the European Liver Fibrosis Group. The methods are discussed briefly below; however, a full description of the study methods is presented in the publications by Rosenberg et al. (29) and Parkes et al. (33). Briefly, serum samples were obtained at the time of liver biopsy from 921 patients recruited consecutively with a spectrum of liver diseases (29) at the time of their first investigation for abnormal liver function. Assays of hyaluronic acid (HA), amino-terminal peptide of procollagen III (PIIINP) and tissue inhibitor of matrix metalloproteinase 1 (TIMP-1) were performed on an Immuno-1 autoanalyser using the manufacturer’s reagents and in accordance with the manufacturer’s instructions (Siemens Healthineers, Tarrytown, USA). Results were entered into the ELF algorithm to calculate ELF scores. Liver biopsies were analysed by a single pathologist and staged for liver fibrosis using both the Scheuer and Ishak systems. In the present study the diagnostic performance of the ELF test was calculated and sensitivity, specificity, and areas under receiver operator characteristic curves are reported. Previously we have also reported positive and negative predictive values for the ELF test (29).

501 patients recruited to the original European Liver Fibrosis study at 7 centres in England (1998-2000) were followed up for clinical outcomes through clinical record review. The spread of aetiologies represented is displayed in Table 1. Of these, 457 were included in the data analysis as 44 patients were excluded because they had entered study with a transplanted liver. The main outcome was the performance of the Enhanced Liver Fibrosis test in predicting the incidence of the first liver-related clinical event defined as: (A) Liver-related death - defined as any mention of liver disease in part 1 of the death certificate (where the primary cause of death is recorded). The
performance of the Enhanced Liver Fibrosis test in predicting all-cause mortality was also assessed; or (B) Any episodes of **Decompensated Cirrhosis** after recruitment – including ascites (detected by paracentesis, ultrasound, or on clinical examination), encephalopathy (defined clinically), oesophageal variceal haemorrhage (confirmed by endoscopy), liver transplantation and hepatocellular cancer (diagnosed by one or more space-occupying lesions seen by imaging methods with typical patterns of HCC or by histology).

Participant socio-demographic and clinical characteristics were described using frequency and percentage for categorical variables, Correlation between continuous variables was assessed using Pearson’s correlation. Linear regression was used to generate and fit regression curves and straight lines for ELF with both Ishak and Scheuer staging. The risks and hazard ratios for liver related events were calculated using Cox proportional hazard model after adjusting for age and sex. Kaplan Meier survival curves for survival free of liver related events including complications of portal hypertension, liver cancer, liver transplantation and death were generated using SPSS for patients with ELF scores in the ranges <7.70; 7.7 - 9.79; 9.80 - 11.29 and ≥11.30. Data analyses were conducted using SPSS Inc version 24.0 (College Station TX: StatCorp LP; 2013).

**Results**

The Delphic panel expressed the view that the clinical utility of non-invasive testing (NIT) is to detect fibrosis in patients considered to be at low risk of severe liver disease; and to identify cases of advanced fibrosis and cirrhosis with accuracy amongst patients suspected of having significant fibrosis.
The clinicians’ consensus was that the test sensitivity for detecting minor
degrees of fibrosis should be 80-85% when seeking cases of fibrosis amongst
patients suspected to be at low risk of fibrosis, acknowledging that this
strategy may miss 20% of cases with minimal fibrosis whose disease may
subsequently progress (false negative test). Their view was that these cases
could be followed with repeated testing if this was clinically indicated. Similarly
the panel opted for 80% sensitivity for the detection of cirrhosis, accepting that
up to 20% of patients with either minimal fibrosis or cirrhosis might be missed
They originally agreed to a minimum specificity for cirrhosis of 85% meaning
that up to 15% of patients would be mis-diagnosed as having cirrhosis when
they in fact had lesser degrees of fibrosis (false positive test).
Subsequently, having reviewed the test performance characteristics of ELF
for the detection of cirrhosis (S6) the clinicians requested an additional
threshold with greater specificity for the diagnosis of cirrhosis such that fewer
than 5% of mild-advanced fibrosis (<S6) cases would be mis-classified as
cirrhotic. This threshold coincides with that previously identified by
Lichtinghagen et al (37). The ELF test thresholds that yielded the sensitivities
and specificities for fibrosis detection identified using the Delphic process,
based on the data obtained from the EUROGOLF cohort (29) and are <7.7 for
none-mild fibrosis, 7.7 to <9.8 for moderate fibrosis, 9.8 to <11.3 for severe
fibrosis and ≥11.3 for cirrhosis.

The ability of ELF to distinguish between different binary categorizations of
stages of fibrosis is presented in Table 2. Using Ishak histological staging as
the reference standard, 80% of patients with any detectable liver fibrosis
 (>S0) and 85% with more than minimal fibrosis (>S1) will have an ELF score of greater than 7.70. In the mid-range between 7.71 and 9.79 the ELF test can be used to classify patients as having mild or moderate fibrosis (>S2) with 80% sensitivity. The ≥9.8 threshold has a specificity for the detection of advanced fibrosis (>S4) of 90%. In relation to cirrhosis, 76% of patients with cirrhosis (S6) on liver biopsy will have an ELF score ≥9.8, however the specificity for the detection of cirrhosis using this cut-off is only 87%. The higher threshold of ≥11.3 has 97% specificity. Thus, only 3% of patients will have a false positive diagnosis of cirrhosis.

However, the loss of sensitivity consequent on using the threshold of 11.3 results in the risk of 62% of patients with cirrhosis being mis-classified because they have an ELF score that falls below this threshold.

Examination of the relationship between ELF scores and histological stages within the categorical boundaries established by the thresholds reveals that there is a positive correlation between rising ELF score and increasing Ishak or Scheuer histology stage (Pearson’s R – 0.598 p<0.001 for Ishak and R² = 0.99527 p<0.001 for Scheuer) (see Figure 1A and B).

Fitting straight lines across these curves revealed that a 0.5 change in ELF corresponded to a change in 1 histological stage in the F1-F3 Scheuer or S2-S5 Ishak range (see Figure 1 C and D). The mean change in ELF per Ishak stage between S1 and S6 was 0.52 and the mean change in ELF per Scheuer stage between F1 and F4 was 0.53. An ELF score of less than 7.7 represents no fibrosis while a score of 8.3 correlates with moderate fibrosis, and the
The difference between scores of 8.7 and 9.4 corresponds to the difference between moderate and advanced histological stages of liver fibrosis.

To evaluate the prognostic performance of the ELF test using the thresholds identified for histological staging, ELF scores were compared to the liver related events for each patient. Participants were sorted according to their liver related event status at (i) overall (ii) 5 years follow up (iii) 6 years follow up, according to one of four discrete ELF score groupings (<7.7; 7.7 to <9.8; 9.8 to <11.3; ≥11.3) (Table 3). The risk of developing a liver related event (LRE) and the hazard ratios of these outcomes adjusted for age and sex were calculated using Cox proportional hazards model and are presented in Table 4. An ELF score of 9.8-11.3 is associated with a 10 fold increase in the relative risk of a liver related event in the next 6 years compared to an ELF score of 7.7-9.8. A patient with an ELF score between 9.8 and 11.3 has a 38% chance of having a liver related event within 6 years and is 21 times more likely to experience such an event compared to a patient with little or no liver fibrosis (ELF score <7.7). At 5 years a patient with a score of >11.3 has a 57% chance of a LRE and is 70 times more likely to have a LRE than a person with score of less than 7.7.

The Kaplan-Meier Survival curves for survival without a liver related event for each of the diagnostic thresholds are presented in figure 2. The separation of the curves for 7.7-9.79, 9.8-11.3 and >11.3 each reaches statistical significance.
Discussion

The principle aim of this work is to place in the public domain the evidence that supports the use of the manufacturer’s thresholds for the interpretation of ELF test results. An additional threshold has been described for more precise identification of patients with cirrhosis. Furthermore we have taken the opportunity to determine the prognostic performance of the manufacturer’s published thresholds and the newly derived cirrhosis threshold in the original EUROGOLF cohort over 7 years.

The Enhanced Liver Fibrosis test has been shown to be an accurate non-invasive test for both assessing the severity of fibrosis in chronic liver disease in a wide range of aetiologies (36-39) and predicting the incidence of liver-related death and cirrhosis in patients (33). Previous studies have described similar performance to other biomarker panels and better diagnostic yield than transient elastography (36).

The ELF threshold of 9.8 identifies cirrhotic patients with approximately 80% sensitivity in line with the requirement of the expert panel. However at this threshold 13%, or approximately 1 in 8 of the patients classified as having cirrhosis would in fact have milder fibrosis. Clinicians reported that this threshold resulted in too high a proportion of their patients being classified as potentially cirrhotic and thus potentially requiring investigation and management of portal hypertension and hepatocellular cancer. A higher threshold more specific for cirrhosis was requested. Accordingly the 11.3 threshold was identified, associated with 97% specificity for cirrhosis so that
the number of false positive results is reduced considerably. Patients whose scores lie between 9.8 and 11.3 should be considered to be highly likely to have advanced fibrosis and thus warrant closer monitoring for disease progression according to clinical circumstances. We suggest that clinical judgment should be used in the management of patients with ELF scores in this range. The selection of thresholds should be dictated by circumstance so that the threshold of 9.8 should be used in case-finding strategies seeking cases of cirrhosis while the threshold of 11.3 should be used to confirm the likely presence of cirrhosis. Incidentally this threshold was identified independently by Lichtenhagen at al. (40) from analysis of a separate cohort of patients. Further studies are required to evaluate the utility of this cirrhosis specific threshold of ELF in determining the need for variceal assessment and surveillance for liver cancers.

While the manufacturer’s thresholds will undoubtedly aid clinicians used to converting NIT scores to their histological equivalent we believe that this approach under-values the information contained in ELF scores as a continuous variable. Furthermore non-invasive tests have the advantage over histological staging in that they can provide further diagnostic and prognostic information even after cirrhosis is established. Liver biopsies are often reported as exhibiting histological stages of fibrosis near the boundaries of categorical variables. The near linear correlation between ELF thresholds and the corresponding Ishak stages provides the opportunity to interpret ELF scores falling between the thresholds in the differentiation between degrees of fibrosis within categorical histology stages. These finer distinctions are of particular value when attempting to monitor disease progression or regression.
where NIT such as ELF may prove most informative in tracking response to
treatment. Increasingly the ELF test is being used in this context\(^\text{41, 42}\). In
the mid ranges of fibrosis severity there appears to be a near linear
correlation between ELF score and histological stage. The greater increase in
median ELF scores associated with cirrhosis \((\text{S6})\) can be attributed to the fact
that ELF scores continue to track increasing amounts of fibrosis within the
single histological category of cirrhosis.

The biological validity of the ELF test as a surrogate measure of liver fibrosis
is suggested by the correlation with metabolic changes associated with more
advanced fibrosis \(^\text{43}\) and is further supported by the good correlation
between ELF scores and prognosis for liver related morbidity and mortality.
The prognostic value of the ELF test has been reported in CLD caused by a
range of aetiologies in the original EUROGOLF cohort\(^\text{33}\), as well as in
primary biliary cholangitis\(^\text{32}\) and in primary sclerosing cholangitis\(^\text{34}\). In
each of these studies thresholds have been derived from original data in order
to predict the risk of clinical outcomes. While these thresholds are informative,
the difference between thresholds for fibrosis assessment and prognostication
introduces complexity and confusion. We have used the thresholds identified
for fibrosis assessment to interrogate the original EUROGOLF cohort for
clinical outcomes. These three thresholds define four categories of risk for
liver related outcomes that are statistically separate and carry meaningful
information that can be conveyed to patients and colleagues and used to
guide management decisions.
Whilst the categorical risk stratification based on Cox proportionate hazard models and Kaplan-Meier survival analysis identified statistically separate groups, as with all categorical groupings movement between categories across a single threshold does not reflect the true change in risk for the patient and so changes in ELF score should be interpreted in relation to their correlation with the biological process of fibrosis that underpins risk of morbidity and mortality. Previously we have shown that a unit change in the ELF score correlates with a doubling of the risk of a liver related outcome over 7 years (33). Thus the difference between a score of 8.4 and 9.5 represents a clinically meaningful and important difference even though both scores fall between the 7.7 and 9.8 thresholds denoting moderate fibrosis.

In the 2016 NICE guidelines the value of 10.51 has been recommended for the assessment and management of NAFLD to diagnose patients with advanced fibrosis (44). This value falls in the middle of the 9.8 – 11.3 range with an associated specificity for diagnosing advanced fibrosis >90%. Furthermore, the risk of a liver-related death in 5 years with an ELF score of 10.5 is approximately 25%, and the patient has a risk at least 20 times greater than someone with an ELF score <7.7.

It must be acknowledged that the thresholds defined in this study are based on specific cohort of patients with mixed chronic liver diseases and may not be equally applicable in all categories of chronic liver disease. Sub-group analyses failed to show any statistically significant differences in the
performance of ELF between viral, alcoholic and fatty aetiologies but larger studies in these disorders are warranted. Although this study has not investigated the performance of ELF in different aetiologies it remains a strength of the test that it can be applied to a wide range of chronic liver diseases using the same thresholds.

This study aimed to compare ELF with the reference standard tests of liver histology and clinical outcomes. Participants were not assessed with fibroelastography or with “indirect’ marker assays such as APRI and FIB4 due to variation in the availability of both AST and ALT results but such comparisons would be of interest. Furthermore it would be of interest to analyse changes of ELF over time and their relationship to clinical outcome. It is hoped that by placing these data in the public domain other investigators will be enabled to explore the utility of the ELF test in the investigation, prognostication and management of patients with chronic liver disease.

Acknowledgements:

Prof. William Rosenberg is a National Institute for Health Research (NIHR) Senior Investigator. He is supported by the NIHR Biomedical Research Centre UCLH, the UCL Health Protection Research Unit and the NIHR Diagnostic Evidence Co-operative Leeds.

This work was supported by the NIHR Wellcome Trust Southampton Clinical Research Facility.
References

1. WHO European Health for All Database. WHO. 2009.
### Table 1: Demographics of the EUROGOLF cohort

<table>
<thead>
<tr>
<th>Clinical Condition</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatitis C Virus (HCV) Infection</td>
<td>54</td>
<td>143</td>
<td>197</td>
</tr>
<tr>
<td>Hepatitis B Virus (HBV) Infection</td>
<td>2</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Non-alcoholic Fatty Liver Disease (NAFLD)</td>
<td>8</td>
<td>36</td>
<td>44</td>
</tr>
<tr>
<td>Alcoholic Liver Disease (ALD)</td>
<td>24</td>
<td>62</td>
<td>86</td>
</tr>
<tr>
<td>Autoimmune Hepatitis (AIH)</td>
<td>20</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>Hereditary Hemochromatosis (HH)</td>
<td>5</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Primary Biliary Cholangitis (PBC)</td>
<td>16</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Primary Sclerosing Cholangitis (PSC)</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>*Other</td>
<td>21</td>
<td>18</td>
<td>39</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>152</td>
<td>305</td>
<td>457</td>
</tr>
</tbody>
</table>

*Includes cases of chronic liver disease of unknown aetiology and cases in which no diagnosis was made in the investigation of abnormal liver function tests.

*Table 1 Legend: 457 patients recruited to EUROGOLF in the United Kingdom were followed up for clinical outcomes over 7 years following enrolment.*
Table 2: The relationship between the ELF Score and liver fibrosis severity

<table>
<thead>
<tr>
<th>ELF Score</th>
<th>Ishak Histological Staging</th>
<th>Fibrosis Severity</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.7</td>
<td>F0 v 123456</td>
<td>Minimal</td>
<td>80</td>
<td>42</td>
</tr>
<tr>
<td>7.7</td>
<td>F01 v 23456</td>
<td>Mild</td>
<td>85</td>
<td>38</td>
</tr>
<tr>
<td>8.0</td>
<td>F01 v 23456</td>
<td>Mild Fibrosis</td>
<td>80</td>
<td>59</td>
</tr>
<tr>
<td>8.3</td>
<td>F012 v 3456</td>
<td>Moderate Fibrosis</td>
<td>80</td>
<td>60</td>
</tr>
<tr>
<td>8.7</td>
<td>F0123 v 456</td>
<td>Moderate Fibrosis</td>
<td>80</td>
<td>72</td>
</tr>
<tr>
<td>9.0</td>
<td>F0123 v 456</td>
<td>Moderate Fibrosis</td>
<td>73</td>
<td>80</td>
</tr>
<tr>
<td>9.4</td>
<td>F01234 v 56</td>
<td>Advanced Fibrosis</td>
<td>73</td>
<td>84</td>
</tr>
<tr>
<td>9.8</td>
<td>F01234 v 56</td>
<td>Advanced Fibrosis</td>
<td>65</td>
<td>90</td>
</tr>
<tr>
<td>≥ 9.8</td>
<td>F012345 v 6</td>
<td>Cirrhosis</td>
<td>76</td>
<td>87</td>
</tr>
<tr>
<td>≥ 11.3</td>
<td>F012345 v 6</td>
<td>Cirrhosis</td>
<td>38</td>
<td>97</td>
</tr>
</tbody>
</table>

Table 2 Legend: ELF scores corresponding to Ishak fibrosis stage groupings as binary outcomes are presented with associated histological descriptors and the sensitivity and specificity for detection of the fibrosis stage.
### Table 3: Clinical Correlation of ELF Score to Liver related outcome

#### (i) 5 years follow-up

<table>
<thead>
<tr>
<th>ELF</th>
<th>Liver related outcome at 5 years follow up number (%)</th>
<th>No liver related outcomes number (%)</th>
<th>Total number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7.7</td>
<td>1 (1.1)</td>
<td>88 (98.9)</td>
<td>89 (100)</td>
</tr>
<tr>
<td>7.70 - 9.79</td>
<td>9 (4.1)</td>
<td>208 (95.9)</td>
<td>217 (100)</td>
</tr>
<tr>
<td>9.80 - 11.29</td>
<td>13 (23.6)</td>
<td>42 (76.4)</td>
<td>55 (100)</td>
</tr>
<tr>
<td>≥11.3</td>
<td>21 (56.8)</td>
<td>16 (43.2)</td>
<td>37 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>44 (11.1)</td>
<td>354 (88.9)</td>
<td>398 (100)</td>
</tr>
</tbody>
</table>

#### (ii) 6 years follow-up

<table>
<thead>
<tr>
<th>ELF</th>
<th>Liver related outcome at 6 years follow up number (%)</th>
<th>No liver related outcomes number (%)</th>
<th>Total number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7.7</td>
<td>2 (2.5)</td>
<td>79 (97.5)</td>
<td>81 (100)</td>
</tr>
<tr>
<td>7.70 - 9.79</td>
<td>11 (5.6)</td>
<td>187 (94.4)</td>
<td>198 (100)</td>
</tr>
<tr>
<td>9.80 - 11.29</td>
<td>20 (37.7)</td>
<td>33 (62.3)</td>
<td>53 (100)</td>
</tr>
<tr>
<td>≥11.3</td>
<td>21 (56.8)</td>
<td>16 (43.2)</td>
<td>37 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>54 (14.6)</td>
<td>315 (85.4)</td>
<td>369 (100)</td>
</tr>
</tbody>
</table>

#### (iii) Overall (7 years follow-up)

<table>
<thead>
<tr>
<th>ELF</th>
<th>Liver related outcome number (%)</th>
<th>No liver related outcomes number (%)</th>
<th>Total number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7.7</td>
<td>2 (1.9)</td>
<td>104 (98.1)</td>
<td>106 (100)</td>
</tr>
<tr>
<td>7.70 - 9.79</td>
<td>12 (4.8)</td>
<td>239 (95.2)</td>
<td>251 (100)</td>
</tr>
<tr>
<td>9.80 - 11.29</td>
<td>23 (37.7)</td>
<td>38 (62.3)</td>
<td>61 (100)</td>
</tr>
<tr>
<td>≥11.3</td>
<td>24 (61.5)</td>
<td>15 (38.5)</td>
<td>39 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>61 (13.3)</td>
<td>396 (86.7)</td>
<td>457 (100)</td>
</tr>
</tbody>
</table>

*Table 3 Legend:* The number of liver related outcomes (including complications of portal hypertension, liver cancer, liver transplantation and death) are presented (i) at 5 years; (ii) at 6 years and (iii) for the duration of follow-up to 7 years;
<table>
<thead>
<tr>
<th></th>
<th>Risk % (95% CI)</th>
<th>Hazard Ratio (95% CI) Cox Proportional Hazard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver related outcomes 5 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7.7</td>
<td>1.1 (0.2, 6.1)</td>
<td>1.0</td>
</tr>
<tr>
<td>7.70 - 9.79</td>
<td>4.1 (2.2, 7.7)</td>
<td>3.53 (0.45, 28.12)</td>
</tr>
<tr>
<td>9.80 - 11.29</td>
<td>23.6 (14.4, 36.3)</td>
<td>21.00 (2.68, 164.65)</td>
</tr>
<tr>
<td>≥11.3</td>
<td>56.8 (40.9, 71.3)</td>
<td>71.04 (9.4, 536.7)</td>
</tr>
<tr>
<td>Liver related outcomes 6 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7.7</td>
<td>2.5 (0.7, 8.6)</td>
<td>1.00</td>
</tr>
<tr>
<td>7.70 - 9.79</td>
<td>5.6 (3.1, 9.7)</td>
<td>2.07 (0.44, 9.67)</td>
</tr>
<tr>
<td>9.80 - 11.29</td>
<td>37.7 (25.9, 51.2)</td>
<td>21.03 (4.60, 99.19)</td>
</tr>
<tr>
<td>≥11.3</td>
<td>56.8 (40.9, 71.3)</td>
<td>47.37 (9.75, 230.17)</td>
</tr>
</tbody>
</table>

Table 4 Legend: The risk and hazard ratios of a liver related event associated with ELF scores in specified ranges are presented along with the associated 95% confidence intervals. Hazard ratios are calculated using Cox proportional hazard model after adjusting for age and sex.
Figure Legends:

Figure 1A: The relationship between ELF score and histological stage of liver fibrosis using the Ishak staging system.

Figure 1B: The relationship between ELF score and histological stage of liver fibrosis using the Scheuer staging system.

Figure 1C: Linear regression was used to fit a line to the plot of median ELF scores against Ishak histological stages. The line of best fit was used to determine the relationship between ELF and fibrosis stage. An increase of approximately 0.5 in ELF corresponds to an increase in fibrosis of 1 Ishak stage in the mid range from S1 to S5. The greater increase in median ELF scores associated with cirrhosis (S6) can be attributed to the fact that ELF scores continue to track increasing amounts of fibrosis within the single histological category of cirrhosis.

Figure 1D: Linear regression was used to fit a line to the plot of median ELF scores against Scheuer histological stages. The line of best fit was used to determine the relationship between ELF and fibrosis stage. An increase of approximately 0.5 in ELF corresponds to an increase in fibrosis of 1 Scheuer stage across the mid range from F1 to F4. The greater increase in median ELF scores associated with cirrhosis (F4) can be attributed to the fact that ELF scores continue to track cirrhosis within the single histological category of cirrhosis.

Figure 2: Kaplan Meier survival curves for survival free of liver related events including complications of portal hypertension, liver cancer, liver transplantation and death for patients with ELF scores in the ranges <7.70; 7.7 - 9.79; 9.80 - 11.29 and ≥11.30
Figure 2

[Graph showing Time to liver related events with different ELF IFU thresholds]