Morling, Joanne R. and Fallowfield, Jonathan A. and Guha, Indra Neil and Nee, Lisa D. and Glancy, Stephen and Williamson, Rachel M. and Robertson, Christine M. and Strachan, Mark W.J. and Price, Jackie F. (2014) Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. Journal of Hepatology, 60 (2). pp. 384-391. ISSN 1600-0641

Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh Type 2 Diabetes Study

[short running title: Non-invasive biomarkers of hepatic fibrosis in type 2 diabetes.]

Joanne R Morling[1], Jonathan A Fallowfield[2], Indra N Guha[3], Lisa D Nee[4], Stephen Glancy[4], Rachel M Williamson[5], Christine M Robertson[1], Mark WJ Strachan[5], Jackie F Price[1]

on behalf of the Edinburgh Type 2 Diabetes Study investigators

1. Centre for Population Health Sciences, The University of Edinburgh, Edinburgh, EH8 9AG
2. MRC Centre for Inflammation Research, University of Edinburgh
3. NIHR Nottingham Digestive Diseases Biomedical Research Unit, University of Nottingham
4. Radiology Department, Western General Hospital, Edinburgh
5. Metabolic Unit, Western General Hospital, Edinburgh

Correspondence to:

Dr Joanne Morling

Centre for Population Health Sciences
University of Edinburgh
Medical School, Teviot Place
Edinburgh EH8 9AG

Tel: 0131 650 3244

Fax : 0131 650 6909

Email: j.morling@ed.ac.uk

Word count: 5,170

Figures: 0

Tables: 4

Online supplementary figures: 3

Abbreviations: NAFLD non-alcoholic fatty liver disease; NAFL non-alcoholic fatty liver; NASH non-alcoholic steatohepatitis; CLD chronic liver disease; NFS NAFLD Fibrosis Score; BMI body mass index; ET2DS Edinburgh Type 2 Diabetes Study; LDR Lothian Diabetes Register; USS ultrasound scanning; TE transient elastography; ALT alanine aminotransferase; AST aspartate aminotransferase; HA hyaluronic acid; P3NP aminoterminal peptide of pro-collagen III; TIMP1 tissue inhibitor of matrix metalloproteinase 1; SCD skin capsule distance; LSM liver stiffness measure; CVH chronic viral hepatitis; PBC primary biliary cirrhosis; AST/ALT ratio alanine aminotransferase to aspartate aminotransferase ratio; APRI aspartate to platelet ratio index; ELF European Liver Fibrosis panel; FIB4 Fibrosis-4 Score; kPa kilopascals; NPV negative predictive value; spec. specificity.

Conflict of interest: None.

Financial support: The ET2DS was funded by a grant from the UK Medical Research Council and the Liver Substudy was supported by a grant from Pfizer. J.R.M. was supported by a Diabetes UK Clinical Research Fellowship. J.A.F is supported by a Clinician Scientist Fellowship from the Academy of Medical Sciences/ The Health Foundation.

Abstract

Background and Aims

It is difficult to determine the different stages of non-alcoholic fatty liver disease without the use of invasive liver biopsy. In this study we investigated five non-invasive biomarkers used previously to detect hepatic fibrosis and determined the level of agreement between them in order to inform future research.

Methods

In the Edinburgh Type 2 Diabetes Study, a population-based cohort aged 60-74 years with type 2 diabetes, 831 participants underwent ultrasound assessment for fatty liver and had serum aspartate aminotransferase to alanine aminotransferase ratio (AST/ALT), aspartate to platelet ratio index (APRI), European Liver Fibrosis panel (ELF), Fibrosis-4 Score (FIB4) and liver stiffness measurement (LSM) measured.

Results

Literature based cut-offs yielded marked differences in the proportions of the cohort with probable liver fibrosis in the full cohort. Agreement between the top 5% of the distribution for each biomarker pair was poor. APRI and FIB4 had the best positive agreement at 76.4%, but agreement for all of the other serum biomarker pairs was between 18% and 34%. Agreement with LSM was poor (9% to 16%).

Conclusions

We found poor correlation between the five biomarkers of liver fibrosis studied. Using the top 5% of each biomarker resulted in good agreement on the absence of advanced liver disease but poor agreement on the presence of advanced disease. Further work is required to validate these markers against liver biopsy and to

determine their predictive value for clinical liver-related endpoints, in a range of different low and high risk population groups.

Word count: 238

Key words: hepatic fibrosis, non-invasive biomarker tests, non-alcoholic fatty liver disease, type 2 diabetes mellitus.

Introduction

Liver dysfunction in people with type 2 diabetes mellitus is thought to be mainly caused by non-alcoholic fatty liver disease (NAFLD). The earliest stage of NAFLD is simple steatosis but this can progress to non-alcoholic steatohepatitis (NASH) and ultimately to hepatic fibrosis, cirrhosis and the long term complications of chronic liver disease (CLD) such as hepatocellular carcinoma. The prevalence of NAFLD is thought to be higher in type 2 diabetes than in the general population[1-4]. Research focusing on the identification of fatty liver using ultrasound suggests a prevalence of around 34% in the general population[1] ; in type 2 diabetes our own group found the prevalence to be 42.6%[2] and this figure may rise to 70% in more selected sub-populations of diabetes [3, 4]. The prevalence of NASH and NASH-related fibrosis is much harder to determine as currently the only widely accepted diagnostic method is liver biopsy. However, it is difficult to justify performing liver biopsy to determine the severity of liver disease in community based subjects, including volunteers in research settings for two key reasons i) there is considerable variability in sampling and histopathological interpretation due to the small volume of tissue sampled (typically 0.002% of the liver)[5] and subjective semi-quantitative scoring systems[6-8], and ii) biopsy is associated with an adverse outcome profile including pain, bleeding and rarely death[9-12]. Thus, there is considerable interest in the adoption of validated non-invasive markers of fibrosis into clinical practice. In the few biopsy studies of populations with type 2 diabetes, the prevalence of advanced fibrosis in those with NAFLD was 7-12%[13-15].

Non-invasive markers have been extensively validated in secondary care for the diagnosis of hepatic fibrosis either for specific underlying pathologies (e.g. the NAFLD Fibrosis Score, NFS) or with varying disease specific cut-offs. There are three broad groups of biomarkers: single markers, combination marker panels and imaging. Increasing numbers of scales and scores are being developed, with most studies reporting acceptable diagnostic accuracy (AUC>0.7) for individual methods in diagnosing the presence of hepatic fibrosis in NAFLD. However, their reliability and utility in identifying undiagnosed liver fibrosis in wider

clinical practice and in research settings is yet to be determined given the limited studies in primary care[16, 17]. Our group has previously shown[18] that the utility of many simple marker panels (BAAT score, BARD score, NFS) is limited in a population with type 2 diabetes by the inclusion of age, body mass index (BMI) and diabetes and led to over-estimation of the prevalence of fibrosis and high levels of indeterminate results.

In this study we investigated five biomarkers used previously to detect hepatic fibrosis in clinical populations with NAFLD. We aimed to determine the level of agreement between these biomarkers, in a large, representative, well-phenotyped population of people with type 2 diabetes mellitus (the Edinburgh Type 2 Diabetes Study, ET2DS).

Patients and Methods

Study population

Full methods of the ET2DS have been published previously[19]. In brief, patients aged 60 to 75 years were selected at random from the Lothian Diabetes Register (LDR), a comprehensive register of patients with diabetes living in Lothian, Scotland. 1066 patients were recruited and attended a baseline clinic for physical examination. Study recruits have been shown previously to be largely representative of all those randomly selected to participate (n=5454) and therefore of the target population of older men and women with type 2 diabetes living in the general population[20]. Participants who were able and willing (n=939) attended a liver assessment 1 year after baseline, including liver ultrasound scanning (USS)[21]. Subjects who were still living were invited to a further detailed assessment approximately four years after recruitment; these subjects (n=831) form the study population for the current analysis.

Ethical approval was obtained from the Lothian Research Ethics Committee and all subjects gave written informed consent.

Clinical examination and liver assessment

Clinical examination at the year 1 liver assessment and year 4 follow-up was similar to that performed in earlier phases of the study, described in detail previously[19]. In brief, patients underwent physical examination (including height and weight measurements); venepuncture; self-administered questionnaire (including alcohol consumption) and liver imaging (including USS and transient elastography (TE)). Plasma glucose, HbA1c, platelets and liver enzymes (including alanine aminotransferase (ALT), aspartate aminotransferase (AST)) and albumin were measured on a fasting blood sample using a Vitros Fusion chemistry system (Ortho Clinical Diagnostics, Bucks, UK). European Liver Fibrosis (ELF) panel - comprising hyaluronic acid (HA), aminoterminal peptide of pro-collagen III (P3NP) and tissue inhibitor of

matrix metalloproteinase 1 (TIMP1) - was measured using the ADVIA Centaur immunoassay system (Siemens Healthcare Diagnostics Inc, New York, USA) on serum stored at -80^{o}C. USS was performed using a Sonoline Elegra Ultrasound Imaging System (Sieman's Medical Systems Inc, Washington, USA), software version 6, using a 3.5 MHz transducer. A phantom (411 LE 0.5, GAMMEX rmi Ltd, Nottingham, UK) as described and validated previously using magnetic resonance spectroscopy[21] , hepatic steatosis was graded as present or absent based on standard criteria.

One dimensional TE was performed using a FibroScan (Echosens, Paris, France) machine at the year 4 follow-up visit only. A single operator was formally trained by Echosens personnel prior to commencement of the study. Initial ultrasound assessment allowed measurement of the skin-capsule distance (SCD). For SCDs <2.5cm the M probe was used, for SCDs ≥2.5cm the XL probe was used in accordance with recommended standard Fibroscan operating procedures. The TE probe was placed in an intercostal space overlying the liver with the patient in the supine position. Using ultrasound to guide positioning, an area of the liver that was at least 6cm deep and free from large vessels was selected for investigation. The area measured was between 25mm-65mm below the surface of the skin for the M probe and 35mm-75mm for the XL probe. The operator aimed to obtain ten valid liver stiffness measurements (LSM) with a success rate of at least 60% and IQR <30% of the final (median) result. All scans were undertaken in the fasting state (minimum 4 hours). Every six months the probes were serviced and calibrated.

Any patient with plasma liver enzymes above the upper reference limit, any abnormality on liver USS (including steatosis) or LSM > 8kPa underwent a liver screen including viral serology, alpha-feto protein, ferritin, autoantibodies, immunoglobulins, caeruloplasmin and α1-antitrypsin. In addition, pre-diagnosed liver disease was identified from NHS National Services Scotland, Information Services Division data linkage to SMR01 general and acute inpatient discharge records and from patient self-report questionnaires on prior health conditions. Any liver disease identified from linkage and the patient questionnaire was confirmed using individual patient medical records and patients with confirmed pre-diagnosed liver disease

(chronic viral hepatitis, CVH, haemochromatosis and primary biliary cirrhosis, PBC) were excluded from the final analyses.

Data analysis

The five biomarkers/panels evaluated in this investigation were derived as follows:

- Aspartate aminotransferase to alanine aminotransferase ratio (AST/ALT) calculated as AST(U/L) / ALT(U/L).
- Aspartate to platelet ratio index (APRI) calculated as [AST(U/L) / upper limit normal] / platelets(x10^9/L)]x100[22].
- European Liver Fibrosis panel (ELF) calculated as 2.588 + (ln(HA)*(ln(P3NP)*0.775) + (ln(TIMP1)*0.494)[23].
- Fibrosis-4 Score (FIB4) calculated as [age(years) x AST(U/L)] / [platelets(x10^9/L) x √AST(U/L)][24].
- Liver stiffness measurement in kilopascals (kPa) expressed as the median TE value from at least ten valid measurements.

Absolute change in serum biomarker was defined as the change between the year 1 liver assessment and year 4 follow-up. Serum biomarker change was also defined categorically as: increased (increase of >5% of liver assessment value); decreased (decrease of >5% of liver assessment value); and stayed the same (absolute change within 5% of liver assessment value).

NAFLD was defined as the presence of hepatic steatosis on USS without alcohol excess or use of hepatotoxic medication and a negative liver screen. Alcohol excess was defined according to established criteria as alcohol intake >14 units/week (female) or >21 units/week (male)[25], or participant self-report of current/previous alcohol excess[26]. Use of hepatotoxic medication included the use of (non-topical) glucocorticoids for >2 weeks, isoniazid, methotrexate, amiodarone, or tamoxifen within the 6 months prior

to USS[3, 26]. Clinically significant positive immunology titres were defined as ASmA titre >1:160 or AMA titre >1:40[26, 27].

All patients with data available for APRI, AST/ALT ratio, ELF, and FIB4 were included in the analysis. All continuous variables were assessed for approximation to the normal distribution with APRI and FIB4 showing a skewed distribution. Correlation between biomarkers was analysed after standardisation to Z-scores, and adjusted for age and sex. Cronbach's alpha was used to examine the inter biomarker agreement (using standardised Z-scores). Student's t-test or the Mann-Whitney U test were used to compare means and Chi-squared test to compare proportions.

Validated cut-offs to reliably exclude advanced fibrosis (≥ metavir F3) in NAFLD were determined from the literature aiming to achieve negative predictive values (NPVs) 90-95%: APRI=1.0, specificity (spec.) 89%, NPV 84%[22, 28]; AST/ALT=1.0, spec. 90%, NPV 89%[28]; ELF=10.358, spec. 94%, NPV 90%[23]; FIB4=1.30, spec. 65%, NPV 95%[24, 28]; and LSM=8.7, spec. 83.2%, NPV 94.6%[29]. However, since these threshold levels cannot reliably be extrapolated from the predominantly secondary care settings in which they were validated to a general population setting, we also assessed agreement between biomarkers using the same highest percentile across all biomarker panels. Prior studies suggest that the prevalence of advanced fibrosis in type 2 diabetes patients attending outpatient clinics is at least 2-6% in all patients and 7-12% in those with NAFLD[13-15]. As a result we estimated that the underlying prevalence of significant hepatic fibrosis in the whole ET2DS cohort might be in the region of 5% and around 10% in those with NAFLD. We therefore compared the top 5% (and 10%) of scores for each biomarker for agreement in the entire cohort (and in those with NAFLD respectively). Individual 2x2 tables were calculated for the absence/presence of probable fibrosis (based on percentiles) for each pair of biomarkers. Due to the difficulties interpreting markers of total agreement[30] (e.g. kappa statistics), we calculated positive agreement (agreement on the presence of fibrosis by both biomarkers) and negative agreement (agreement on the absence of fibrosis by both biomarkers)[31].

Correlation between absolute changes in serum biomarkers (i.e. between year 1 and year 4) was analysed after standardisation to Z-scores, and adjusted for age and sex. Categorical change in biomarker was analysed using the Chi-squared test to compare proportions.

Analysis was undertaken on both the whole population without pre-diagnosed liver disease and additionally in those with NAFLD (defined as the presence of hepatic steatosis on USS without alcohol excess, use of hepatotoxic medication or abnormal liver screen).

		Biomarker 1		
		Fibrosis present	Fibrosis absent	Total
Biomarker 2	Fibrosis present	a	b	a+b
	Fibrosis absent	c	d	c+d
	Total	a+c	b+d	a+b+c+d

Positive agreement = 2a/(2a+b+c)

Negative agreement = 2d/(b+c+2d)

Results

Participants

831 participants (78% of the original cohort) attended the 4-year follow-up clinic. There were no significant differences between attenders at baseline and at the 4-year visit (Table 1). Of the subjects that did not attend the 4-year clinic (n=235), nine had withdrawn from the study, 14 were un-contactable, 124 were unable or unwilling to attend and 88 had died. Eleven patients were excluded from the analysis due to pre-diagnosed liver disease.

767 participants had a complete biomarker dataset (excluding LSM) available for analysis. Of these, valid LSM was reported for n=650. Missing plasma and serum markers were missing at random. LSM was missing in 117 participants (15.3%) due to failure to obtain valid readings. Patients with missing LSM were significantly more obese, with poorer glucose control and more severe markers of liver dysfunction (Table 2). There were 282 (36.8%) participants fulfilling the criteria for NAFLD with a complete biomarker dataset, with LSM available on 248.

Biomarker distributions and agreement

Biomarker distributions for the full study population are available in the supplementary material (Supplementary Figure A) and correlations between the biomarkers in Table 3 (and Supplementary Figure B). After adjustment for age and sex, correlation was strong between APRI and FIB4 (r=0.92) but all others were ≤0.5. The inter-item correlation for all five markers was just below acceptable (α=0.67) with minimal improvement with the removal of any marker. Correlations were similar in the NAFLD cohort.

Literature based cut-offs yielded marked differences in the proportions of the cohort with probable liver fibrosis in the full cohort: APRI 0.8%, AST/ALT ratio 22.4%, ELF 7.0%, FIB4 68.3%, and LSM 4.5%; and in the NAFLD subgroup: APRI 0.4%, AST/ALT ratio 16.7%, ELF 4.3%, FIB4 63.8%, and LSM 4.8%.

Agreement between the top 5% of the distribution for each biomarker pair was poor (Table 4). APRI and FIB4 had the best positive agreement at 76.4%, but agreement for all of the other biomarker pairs was between 9% and 32%. When the comparison groups were altered to reflect the top 3% and 7% of the biomarker distributions, then with the exception of APRI and FIB4, the agreement did not alter greatly (agreement 9-26% and 11-38% respectively). When analyses were restricted to the study population with NAFLD, agreement between the top 10% of the distribution for each biomarker pair remained generally poor (agreement for APRI and FIB4 was 68%, all other pairs between 14% and 36%), as did agreement for the top 5% and 15% of the distributions (agreements from 0% to 36% and from 21% to 36% respectively, with exception of APRI/FIB4).

The top 5% (10% in NAFLD) was suggestive of advanced liver dysfunction with clinical data implicating advanced fibrosis/cirrhosis with platelet counts being significantly lower and spleen size being significantly larger in the majority of cases (data not shown).

Negative agreement (agreement on the absence of fibrosis) was more consistent for both the full cohort analysis and the NAFLD subgroup (90-99%) with minimal change with the alternative cut-offs described. 2x2 tables for agreement using the top vigintiles are available in the supplementary material (Supplementary Table A).

Changes in biomarkers

Serum biomarker pairs from the year 1and year 4 follow-up were available for 534 subjects (LSM was not undertaken at year 1). Six patients were excluded for pre-existing liver disease leaving an analysis cohort of 528 subjects. There were 183 (34.7%) participants fulfilling the criteria for NAFLD.

Following a mean follow-up of 3.5 years, mean percentage (sd) changes in serum biomarkers were: AST/ALT ratio -0.04 (0.25), APRI 0.19 (0.37), ELF 0.03 (0.08) and FIB4 0.18 (0.30). The changes in each biomarker (adjusted for age and sex) were poorly correlated (AST/ALT:APRIr=0.23 $p<0.001$,

AST/ALT:ELF r=0.15 p<0.001, AST/ALT:FIB4 r=0.55 p<0.001, APRI:ELF r=0.12 p=0.008, ELF:FIB4 r=0.13 p=0.003) with the exception of APRI and FIB4 (r=0.88, p<0.001). The results were similar when limited to patients with NAFLD (Supplementary Figure C). When classified as increasing, staying the same and decreasing, agreement remained weak (35.0% to 54.7%), again with the exception of APRI and FIB4 (76.3% agreement).

Discussion

This study is the first population based study to compare the distribution of different non-invasive markers of liver fibrosis in patients with type 2 diabetes. Its strength lies in its population based approach allowing the results to be generalised to the wider type 2 diabetes population and not just patients at the more severe end of the diabetes spectrum and/or those with known fatty liver disease attending hospital clinics. We found poor correlation between the five biomarkers of liver fibrosis studied. Using the top vigintile (5%) of each biomarker resulted in excellent agreement on the absence of advanced liver disease but poor agreement on the presence of advanced liver disease.

An interesting finding of this cohort is the likely small numbers of patients with hepatic fibrosis. Based on strict application of validated biomarker cut offs used previously in non-diabetic populations to indicate the presence of fibrosis, the prevalence of significant fibrosis (Metavir F3+) appears to be between 1% and 68%, but is most likely less than 10%. Since NAFLD is increasingly reported as associated with type 2 diabetes, it is perhaps surprising (though also encouraging) to find that the prevalence of advanced liver disease may be low in this perceived high risk group. There are several possible reasons for a low prevalence of fibrosis in our study population. It may be related to specific cohort effects and survival bias, with only 'low risk' adults with type 2 diabetes surviving to the age of 60 years and thereby enabling participation in the study. It may be 'artefact', due to behaviour of the biomarkers in populations with diabetes, such that cut-points which indicate fibrosis in a non-diabetic person do not reflect the same level of fibrosis in a diabetic person. There is little literature on the distribution of biomarkers in exclusively diabetic populations, however, there are no clear biologically plausible reason to expect different thresholds to operate. Thus, it may also reflect a truely low prevalence of fibrosis in people with diabetes. This has not been widely studied previously, but given the strong association between insulin insensitivity and hepatic fibrosis development[32] it may be that the intensive management of diabetic patients with insulin sensitising agents (metformin[33] and peroxisome proliferator-activated receptor-gamma agonists (thiozolidinediones)[34]) is attenuating the development of

fibrosis. The probable low prevalence of fibrosis in this cohort is a limitation when analysing biomarker agreement and, due the well-established influence of disease prevalence on diagnostic test accuracy, it may be that there would be a higher biomarker agreement in populations with higher fibrosis prevalence[35].

In addition, 15% of LSM data was missing. The XL probe is believed to increase the success rates of obtaining ≥10 valid TE readings from 56% to 75%[36] in obese populations. This is consistent with our 85% success in a mixed overweight and obese cohort. Despite the use of the XL probe, TE still appears to be limited by body habitus. Those patients with missing LSM had more severe diabetes profiles and higher alternative biomarkers of hepatic fibrosis. Hence, it was felt necessary to expand the majority of analysis to those patients with a full set of biomarkers excluding LSM in order to avoid biasing the sample.

Overall there was poor correlation between the biomarkers measured in this study. The only exception was APRI and FIB4 ($r=0.92$ in full cohort and 0.93 in NAFLD subgroup, $p<0.001$), which is unsurprising given they have a number of components in common (AST level and platelets). Assessment of inter-item correlation ($\alpha=0.67$) determined that the five biomarkers were not consistent with each other in what they were measuring, with a score of $\alpha=0.70$ being the minimum score usually accepted. However, a minimum of 0.90 is often suggested for clinical practice[37]. This suggests some discrepancy in what the biomarkers are measuring. This seems plausible as AST/ALT, APRI and FIB4 are all similar in their composition, including markers of hepatocellular damage. Alternatively, ELF is measuring markers related to extracellular matrix turnover in fibrosis, and LSM is also examining structural properties of the liver through shear wave transmission.

A higher prevalence of fibrosis in the NAFLD cohort compared to the full cohort was expected. However, using validated cut-offs, our results showed the opposite. This probably reflects the natural history of NAFLD in which fibrotic progression is often associated with steatosis regression and hence this group of advanced liver disease patients might not be captured by our definition of NAFLD. ELF scores and LSM values for the top vigintile were most in keeping with the values one would expect to find from previous

published studies[13-15], however, with no reference standard (biopsy) it is impossible to establish which of the five biomarkers is the most accurate and to comment on the true prevalence of hepatic fibrosis using non-invasive biomarkers.

In addition, using validated cut-offs we found a wide range of fibrosis prevalence results for the different biomarkers. The reasons for the wide discrepancies are probably two-fold. Firstly, as demonstrated by the lack of inter-correlation, there is an inconsistency in what the different biomarkers/panels are measuring. Secondly, published fibrosis cut-offs from clinical validation studies do not appear to translate readily into research or clinical practice. The difficulty with this is that the predictive value of a test (unlike sensitivity and specificity) is influenced by the prevalence of underlying disease. Most validation cohorts have typically comprised a high proportion of patients with advanced liver disease selected from tertiary referral centres. Our study population has a presumed lower prevalence of advanced liver disease and consequently the predictive values are likely to be different. This resulting lower positive predictive value and higher negative predictive value would mean that literature based cut-offs are neither reliable nor directly comparable with one another in different patient cohorts such as ours. Without liver biopsy (the current reference standard) it is not possible to decide which biomarker is 'best suited' for diagnosing significant hepatic fibrosis in a lower prevalence population.

There are numerous other panel markers of hepatic fibrosis available e.g. BARD, BAAT, NFS. These are typically simple scoring systems using easily available plasma results and patient data. Previous work from our group has found that these scores are likely to overestimate the prevalence of hepatic fibrosis in populations similar to ours as they rely heavily on the incorporation of impaired glucose tolerance, age and body mass index. For example, the prevalence of fibrosis using the BARD and BAAT scores was 92.6% and 79.3% respectively, with the NFS predicting 16.4% fibrosis and 66.8% indeterminate[18]. It is therefore necessary to concentrate on the development of hepatic fibrosis markers that are independent of the underlying characteristics of the population under study.

As we observed for the cross-sectional distribution of serum biomarkers, changes in the biomarkers over 3.5 years were also poorly correlated with the exception of the similarly derived biomarkers (APRI and FIB4). Even after categorising change as increasing, staying the same or decreasing, agreement remained weak. This is not surprising given the weak associations in cross-sectional analysis, however this may also reflect the short follow-up period.

A further area which requires clarification is the most appropriate use of hepatic biomarkers in the general/healthy population. Without this information, it is hard to predict whether expected values are likely to differ in discrete populations, such as the elderly or people with diabetes. Normal routine liver function tests vary with age, sex and ethnicity[38, 39], and therefore any fibrosis marker panel including these components might benefit from specific reference ranges that reflect the individual population.

At present, use of these non-invasive biomarkers is limited within low disease prevalence settings to excluding advanced liver disease. Further work is required to validate these markers for the presence of advanced liver disease against the liver biopsy, but more importantly to determine their predictive value for clinically relevant liver-related endpoints, such as hepatocellular carcinoma, oesophageal varices and cardiovascular outcomes, in a range of different low and high risk population groups, most notably community settings. The practical and ethical challenges of large scale liver biopsy in 'normal' patients persist, but emerging techniques, such as magnetic resonance elastography[40], have the potential to become a more acceptable reference standard.

Acknowledgements

The authors would like to thank Echosens (Paris, France) for kindly loaning the Fibroscan imaging equipment.

Funding

The ET2DS was funded by a grant from the UK Medical Research Council and the Liver Substudy was supported by a grant from Pfizer. J.R.M. was supported by a Diabetes UK Clinical Research Fellowship. J.A.F is supported by a Clinician Scientist Fellowship from the Academy of Medical Sciences/ The Health Foundation.

Author contributions

JRM designed the study, collected and analysed data and wrote the manuscript. JAF researched data, contributed to the discussion and reviewed/edited the manuscript. ING researched data and reviewed/edited the manuscript. LDN collected data and reviewed/edited the manuscript. SG designed the study and reviewed/edited the manuscript. RMW designed the study, collected data and reviewed/edited the manuscript. CMR collected data and reviewed/edited the manuscript. MWJS designed the study, contributed to the discussion and reviewed/edited the manuscript. JFP designed the study, contributed to the discussion and reviewed/edited the manuscript.

References

[1] Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al. Prevalence of hepatic steatosis in an urban population in the United States: Impact of ethnicity. Hepatology 2004;40:1387-1395.

[2] Williamson RM, Price JF, Glancy S, Perry E, Nee LD, Hayes PC, et al. Prevalence of and Risk Factors for Hepatic Steatosis and Nonalcoholic Fatty Liver Disease in People With Type 2 Diabetes: the Edinburgh Type 2 Diabetes Study. Diabetes Care 2011;34:1139-1144.

[3] Angulo P. Nonalcoholic fatty liver disease. N Engl J Med 2002;346:1221 - 1231.

[4] Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. Diabetes Care 2007;30:1212-1218.

[5] Bravo AA, Sheth SG, Chopra S. Liver Biopsy. New England Journal of Medicine 2001;344:495-500.

[6] Vuppalanchi R, Unalp A, Van Natta ML, Cummings OW, Sandrasegaran KE, Hameed T, et al. Effects of liver biopsy sample length and number of readings on sampling variability in nonalcoholic Fatty liver disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2009;7:481-486.

[7] Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898 - 1906.

[8] Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE, et al. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. Hepatology 2006;44:874-880.

[9] Gilmore IT, Burroughs A, Murray-Lyon IM, Williams R, Jenkins D, Hopkins A. Indications, methods, and outcomes of percutaneous liver biopsy in England and Wales: an audit by the British Society of Gastroenterology and the Royal College of Physicians of London. Gut 1995;36:437-441.

[10] Padia SA, Baker ME, Schaeffer CJ, Remer EM, Obuchowski NA, Winans C, et al. Safety and efficacy of sonographic-guided random real-time core needle biopsy of the liver. Journal of Clinical Ultrasound 2009;37:138-143.

[11] Myers RP, Fong A, Shaheen AAM. Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. Liver International 2008;28:705-712.

[12] Piccinino F, Sagnelli E, Pasquale G, Giusti G, Battocchia A, Bernardi M, et al. Complications following percutaneous liver biopsy: A multicentre retrospective study on 68 276 biopsies. Journal of Hepatology 1986;2:165-173.

[13] Prashanth M, Ganesh HK, Vima MV, John M, Bandgar T, Joshi SR, et al. Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus.[see comment]. J Assoc Physicians India 2009;57:205-210.

[14] Gupte P, Amarapurkar D, Agal S, Baijal R, Kulshrestha P, Pramanik S, et al. Non-alcoholic steatohepatitis in type 2 diabetes mellitus. Journal of Gastroenterology and Hepatology 2004;19:854-858.

[15] Leite NC, Villela-Nogueira CA, Pannain VLN, Bottino AC, Rezende GFM, Cardoso CRL, et al. Histopathological stages of nonalcoholic fatty liver disease in type 2 diabetes: prevalences and correlated factors. Liver International 2011;31:700-706.

[16] Armstrong MJ, Houlihan DD, Bentham L, Shaw JC, Cramb R, Olliff S, et al. Presence and severity of non-alcoholic fatty liver disease in a large prospective primary care cohort. Journal of Hepatology 2012;56:234-240.

[17] Roulot D, Costes J-L, Buyck J-F, Warzocha U, Gambier N, Czernichow S, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. Gut 2011;60:977-984.

[18] Williamson RM, Price JF, Hayes PC, Glancy S, Frier BM, Johnston GI, et al. Prevalence and markers of advanced liver disease in type 2 diabetes. Qjm 2012;105:425-432.

[19] Price J, Reynolds R, Mitchell R, Williamson R, Fowkes FG, Deary I, et al. The Edinburgh Type 2 Diabetes Study: study protocol. BMC Endocrine Disorders 2008;8:18.

[20] Marioni RE, Strachan MWJ, Reynolds RM, Lowe GDO, Mitchell RJ, Fowkes FGR, et al. Association Between Raised Inflammatory Markers and Cognitive Decline in Elderly People With Type 2 Diabetes: The Edinburgh Type 2 Diabetes Study. Diabetes 2010;59:710-713.

[21] Williamson RM, Perry E, Glancy S, Marshall I, Gray C, Nee LD, et al. The use of ultrasound to diagnose hepatic steatosis in type 2 diabetes: intra- and interobserver variability and comparison with magnetic resonance spectroscopy. Clinical Radiology 2011;66:434-439.

[22] Wai C-T, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003;38:518-526.

[23] Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers.[see comment]. Hepatology 2008;47:455-460.

[24] Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, et al. FIB-4: An inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology 2007;46:32-36.

[25] Royal College of Physicians. Inquiry on alcohol guideliines. London; 2011.

[26] Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: Practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 2012;55:2005-2023.

[27] Alvarez F, Berg PA, Bianchi FB, Bianchi L, Burroughs AK, Cancado EL, et al. International Autoimmune Hepatitis Group Report: review of criteria for diagnosis of autoimmune hepatitis. Journal of Hepatology 1999;31:929-938.

[28] McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut;59:1265-1269.

[29] Wong VW-S, Vergniol J, Wong GL-H, Foucher J, Chan HL-Y, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51:454-462.

[30] Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. the problems of two paradoxes. Journal of Clinical Epidemiology 1990;43:543-549.

[31] Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. Journal of Clinical Epidemiology 1990;43:551-558.

[32] Marchesini G, Brizi M, Morselli-Labate AM, Bianchi G, Bugianesi E, McCullough AJ, et al. Association of nonalcoholic fatty liver disease with insulin resistance. The American Journal of Medicine 1999;107:450-455.

[33] Bugianesi E, Gentilcore E, Manini R, Natale S, Vanni E, Villanova N, et al. A randomized controlled trial of metformin versus vitamin E or prescriptive diet in nonalcoholic fatty liver disease. The American Journal Of Gastroenterology 2005;100:1082-1090.

[34] Zhang F, Kong D, Lu Y, Zheng S. Peroxisome proliferator-activated receptor-γ as a therapeutic target for hepatic fibrosis: from bench to bedside. Cellular and Molecular Life Sciences 2013;70:259-276.

[35] Ray P, Manach YL, Riou B, Houle TT. Statistical Evaluation of a Biomarker. Anesthesiology 2010;112:1023-1040 1010.1097/ALN.1020b1013e3181d47604.

[36] Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology 2012;55:199-208.

[37] Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997;314:572.

[38] Siest G, Schiele F, Galteau M-M, Panek E, Steinmetz J, Fagnani F, et al. Aspartate Aminotransferase and Alanine Aminotransferase Activities in Plasma: Statistical Distributions, Individual Variations, and Reference Values. Clin Chem 1975;21:1077-1087.

[39] Prati D, Taioli E, Zanella A, Torre ED, Butelli S, Del Vecchio E, et al. Updated Definitions of Healthy Ranges for Serum Alanine Aminotransferase Levels. Annals of Internal Medicine 2002;137:1-10.

[40] Yin M, Talwalkar JA, Glaser KJ, Manduca A, Grimm RC, Rossman PJ, et al. Assessment of Hepatic Fibrosis With Magnetic Resonance Elastography. Clinical Gastroenterology and Hepatology 2007;5:1207-1213.e1202.

Tables

Table 1. Characteristics of full ET2DS study population and liver biomarker analysis cohort, at baseline, year 1 and year 4. Values are mean (sd), median (IQR) or % (n).

Variable	Full ET2DS cohort at baseline (n=1066)	Analysis cohort (n=767)	
		Year 1	Year 4
Age, years	67.9 (4.2)	68.7 (4.1)	71.4 (4.2)
Sex, % male	51.3% (547)	52.8% (405)	52.8% (405)
Fasting glucose, mmol/L	7.56 (2.1)	6.89 (2.2)	7.75 (2.9)
HbA1c, %	7.40 (1.1)	7.19 (1.1)	7.34 (1.2)
HbA1c, mmol/mol	57.3 (12.2)	55.1 (11.6)	56.7 (13.1)
Duration of diabetes, years	6.71 (3.9-11.3)	7.00 (4.0-12.0)	7.10 (7.6-14.9)
Diabetes treatment			
Diet alone, %	19.3% (193)	19.7% (144)	13.9% (106)
OAHA alone, %	63.4% (652)	65.5% (479)	65.3% (498)
Insulin therapy (+/- OAHA), %	17.3% (178)	14.8% (108)	20.8% (159)
BMI, kg/m^2	31.4 (5.7)	31.2 (5.6)*	31.3 (5.8)
Total cholesterol, mmol/L	4.31 (0.9)	4.16 (0.8)	4.25 (0.9)
HDL cholesterol, mmol/L	1.29 (0.4)	1.23 (0.3)	1.37 (0.4)
Systolic BP,mmHg	133.3 (16.4)	138.3 (18.3)	131.4 (17.9)
Diastolic BP,mmHg	69.1 (9.0)	74.6 (9.4)	69.1 (9.2)
Smoking, % never	38.2% (407)	39.1% (300)*	39.1% (300)
Platelets, x10^9/L	257.7 (68.9)	259.8 (67.7)*	229.8 (66.8)
Steatosis present, %	-	55.1% (423)	50.1% (384)
ALT, U/L	-	33.9 (12.8)	36.0 (11.8)
AST/ALT ratio	-	0.94 (0.3)	0.85 (0.2)
APRI	-	0.24 (0.2-0.3)	0.28 (0.2-0.4)
ELF	-	8.89 (0.8)	9.12 (0.8)
FIB4	-	1.37 (1.1-1.8)	1.56 (1.2-2.0)
LSM, kPa	-	-	5.11 (2.6)

*Measured at baseline visit

ALT alanine aminotransferase; APRI aspartate aminotransferase-platelet ratio index; AST aspartate aminotransferase; BMI body mass index; BP blood pressure; ELF European Liver Fibrosis panel; FIB4

Fibrosis-4 Score; HbA1c glycosylated haemoglobin; LSM liver stiffness measurement; OAHA oral anti-hyperglycaemic agent.

Table 2. Association of risk factors for missing liver stiffness measures. Values are mean (sd), median (IQR) or % (n).

	LSM missing N=119	LSM present N=648	p
Age, years	71.2 (4.3)	71.4 (4.1)	NS
Sex, male	37.0% (44)	55.7% (361)	<0.001
Body mass index, kg/m^2	35.2 (7.7)	30.6 (5.1)	<0.001
Skin capsule distance, cm	2.83 (0.9)	2.32 (0.5)	<0.001
Fibroscan probe, 'M'	42.0% (50)	65.6% (425)	<0.001
Glucose, mmol/L	8.21 (3.4)	7.66 (2.8)	NS
HbA1c, %	7.54 (1.3)	7.30 (1.2)	0.038
HbA1c, mmol/mol	59.0 (14.2)	56.3 (12.9)	
ALT, U/L	36.6 (13.2)	35.8 (11.5)	NS
AST, U/L	31.3 (12.2)	29.3 (9.9)	NS
GGT, U/L	19.0 (12.0-48.5)	16.0 (9.5-27.0)	0.001
AST/ALT ratio	0.88 (0.3)	0.85 (0.2)	NS
APRI	0.30 (0.21-0.44)	0.28 (0.21-0.37)	NS
ELF	9.37 (1.0)	9.08 (0.8)	0.002
FIB4	1.65 (1.21-2.30)	1.55 (1.22-1.96)	NS

ALT alanine aminotransferase; APRI aspartate aminotransferase-platelet ratio index; AST aspartate aminotransferase; ELF European Liver Fibrosis panel; FIB4 Fibrosis-4 Score; GGT gamma-glutamyl transferase; HbA1c glycosylated haemoglobin.

Table 3. Correlations between biomarkers. Values are correlation coefficients, adjusted for age and sex

Table 3A. Full cohort (n=767)

APRI	r=0.30			
ELF	r=0.24	r=0.31		
FIB4	r=0.53	r=0.92	r=0.30	
LSM[a]	r=0.15	r=0.20	r=0.25	r=0.16
	AST/ALT	**APRI**	**ELF**	**FIB4**

Table 3B. NAFL cohort (n=282)

APRI	r=0.32			
ELF	r=0.18	r=0.32		
FIB4	r=0.53	r=0.93	r=0.29	
LSM[b]	r=0.15	r=0.29	r=0.26	r=0.29
	AST/ALT	**APRI**	**ELF**	**FIB4**

[a] n=648

[b] n=248

All $p<0.001$

APRI aspartate aminotransferase-platelet ratio index; AST/ALT aspartate aminotransferase-alanine aminotransferase ratio; ELF European Liver Fibrosis panel; FIB4 Fibrosis-4 Score; LSM liver stiffness measure.

Table 4. Agreement between biomarker pairs for the top 5 or 10% of values, in the full cohort and NAFL subgroup.

Table 4A. Positive agreement (presence of fibrosis) in the top 5% of the full cohort.

APRI	18.4%			
ELF	18.4%	31.6%		
FIB4	34.2%	76.3%	34.2%	
LSM	9.5%	12.7%	15.9%	12.7%
	AST/ALT	**APRI**	**ELF**	**FIB4**

Table 4B. Negative agreement (absence of fibrosis) in the bottom 95% of the full cohort.

APRI	95.7%			
ELF	95.7%	96.4%		
FIB4	96.6%	98.8%	91.8%	
LSM	95.4%	95.5%	95.7%	95.5%
	AST/ALT	**APRI**	**ELF**	**FIB4**

Table 4C. Positive agreement (presence of fibrosis) in the top 10% of the NAFL subgroup.

APRI	28.6%			
ELF	14.3%	28.6%		
FIB4	35.7%	67.9%	32.1%	
LSM	29.2%	20.8%	25.0%	25.0%
	AST/ALT	**APRI**	**ELF**	**FIB4**

Table 4D. Negative agreement (absence of fibrosis) in the bottom 90% of the NAFL subgroup.

APRI	92.1%			
ELF	90.6%	92.1%		
FIB4	92.9%	96.5%	92.5%	
LSM	92.4%	91.5%	9.9%	91.9%
	AST/ALT	**APRI**	**ELF**	**FIB4**

APRI aspartate aminotransferase-platelet ratio index; AST/ALT aspartate aminotransferase-alanine aminotransferase ratio; ELF European Liver Fibrosis panel; FIB4 Fibrosis-4 Score; LSM liver stiffness measure.