Published in final edited form as:

Hepatology. 2017 June ; 65(6): 1891–1903. doi:10.1002/hep.29090.

DUAL PHOTON MICROSCOPY BASED QUANTITATION OF FIBROSIS-RELATED PARAMETERS (Q-FP) TO MODEL DISEASE PROGRESSION IN STEATOHEPATITIS

Yan Wang, MD, PhD¹, Robert Vincent, MD, PhD², Jinlian Yang, MSc¹, Amon Asgharpour, MD², Xieer Liang, MD¹, Michael O. Idowu, MD³, Melissa J. Contos, MD³, Kalyani Daitya, MSc², Mohammed S. Siddiqui, MD², Faridoddin Mirshahi, MSc², and Arun J. Sanyal, MBBS, MD²

¹Dept. of Infectious Diseases & Hepatology Unit, Southern Medical University Nanfang Hospital, GZ China

²Div. of Gastroenterology, Hepatology and Nutrition, Dept. of Internal Medicine, VCU School of Medicine, Richmond, VA

³Div. of Surgical Pathology, Dept. of Pathology, VCU School of Medicine, Richmond, VA

Abstract

There is a need for further refinement of current histological systems for assessment of hepatic fibrosis in nonalcoholic fatty liver disease (NAFLD). This study evaluated hepatic fibrosis in NAFLD using dual photon microscopy based quantitation of fibrosis-related parameters (q-FPs). Fifty (test cohort) and 42 (validation cohort) subjects with NAFLD and the full spectrum of fibrosis were studied. Q-FPs were measured in specific predefined regions of interest (general, vessel, perisinusoid and vascular septa). Seventy q-FPs had inter- and intra-observer concordance 0.8 and were related to the NASH CRN fibrosis staging. Of these, 16 q-FPs with the strongest correlations ($p < 0.001$ for all) were entered in a principal component analysis model (Odds ratio [OR] 7.8, $p < 0.001$), which separated any stage of fibrosis versus no fibrosis, and cirrhosis versus earlier stages with the areas under receiver-operating-characteristic curves of 0.88 and 0.93 ($p = 0.01$ for both), respectively. In an independent multivariable analysis, four q-FPs including the number of collagen strands (OR 8.5, $p = 0.004$), strand length (OR 12.0, $p = 0.02$), strand eccentricity (OR 8.3, $p = 0.004$), and strand solidity (OR 8.0, $p = 0.003$) were independently associated with fibrosis stages and were used to model fibrosis along a continuous linear scale using Desirability functions; this linear scale of fibrosis measurement was also related to fibrosis stage ($p < 0.0001$). The robustness of both the multivariable model and the linear scale of measurement was confirmed in the validation cohort.

Conclusion—Q-FPs model provides an accurate reproducible method to evaluate fibrosis in NAFLD along a quantitative and continuous scale.

Primary author: Yan Wang, MD, PhD, Professor of Medicine and Bioengineering, Dept. of Infectious Diseases & Hepatology Unit, Southern Medical University Nanfang Hospital, No 1838 Guangzhou Dadao Bei Avenue, GZ China 510515, yanwang@smu.edu.cn, +8620 6164 7396

Senior author: Arun J. Sanyal, MBBS, MD, Professor of Medicine, Physiology and Molecular Pathology, MCV Box 980341, Richmond, VA 23298-0341, arun.sanyal@vcuhealth.org, **Phone:** (804) 828 6314, **Fax:** (804) 828 2992

Keywords

Nonalcoholic fatty liver disease; nonalcoholic steatohepatitis; cirrhosis; fibrosis; histology; disease progression; quantitative assessment; fibrosis-related parameter

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is a major cause of liver-related morbidity and mortality and is expected to surpass hepatitis C virus as an etiology for end stage liver disease requiring liver transplantation.¹ NAFLD may manifest as nonalcoholic fatty liver (NAFL) or nonalcoholic steatohepatitis (NASH). NASH progresses to cirrhosis more commonly than NAFL and is therefore the target for therapeutic development.²

A principal objective in drug development for NASH is the prevention of progression to cirrhosis.³ While the architectural remodeling of the liver during progression to cirrhosis includes many components, the fibrosis stage is a robust readout of such progression and is used for histological assessment of disease progression.^{4, 5} The NIDDK NASH CRN system is the most widely used system of fibrosis assessment and classifies the disease into four stages as follows: (1 a–c) based on acinar zone 3 perisinusoidal (1a, 1b) or portal/periportal distribution (1c), (2) zone 3 plus portal/periportal fibrosis, (3) as above with bridging fibrosis and (4) cirrhosis.⁶ While histological assessment of NASH has multiple strengths,⁷ it also has some limitations such as sampling error related to biopsy length and intra- and inter-observer variability in assessment.⁸ Furthermore, such systems convert the continuous range of fibrosis into an ordinal system introducing potential for error at the interface of different stages. The limited dynamic scale of measurement of current staging systems further hinders the ability to identify modest changes in fibrosis.

There are also logistical challenges with using histology based assessment of NAFLD in community clinical practice. This has fueled efforts to identify all aspects of the disease noninvasively. Several circulating and imaging based fibrosis markers are currently in development along with kinetic markers of fibrogenesis.^{9–13} The current development path for such noninvasive methods are based on traditional histological assessment as the gold standard. From a methodological point of view, it is not possible to outperform the gold standard and such development are anchored with the “limitations” of the histological gold standards. There is therefore a need for further refinements in histological assessment of NASH and hepatic fibrosis.

Recently, dual photon microscopy has been used for assessment of hepatic fibrosis.^{14, 15} This method identifies individual collagen fibrils in unstained paraffin-embedded tissue sections and quantifies their physical attributes, such as number, length, diameter, orientation, contour and cross-linkage with each other, within the context of hepatic architecture. It is thus technically quite different from other quantitative methods, such as collagen morphometry, which measures the area stained by collagen-specific stains e.g. Sirius Red and reads it out as a function of the total area measured (collagen proportional area).¹⁶ Proof-of-concept studies have demonstrated that dual photon microscopy-based quantifiable fibrosis-related parameters (q-FPs) can be used to evaluate fibrosis in a rat

model of fibrosis and in chronic hepatitis B.¹⁴ However, it has not been evaluated for NASH which is dominated by perisinusoidal fibrosis and formation of central-portal and central-central vein bridges and thus, is quite different than the mainly portal pattern of fibrosis seen with chronic viral hepatitis.¹⁷

The overall objective of this study was to perform a “proof-of-concept” study to evaluate and validate the ability of q-FP based analysis to quantitatively assess the hepatic fibrosis in NASH. The specific aims were to: (1) describe the spectrum of q-FPs that could be assessed in liver biopsies from humans with varying stages of NAFLD, (2) quantitate each measured q-FP and evaluate the test-retest repeat variability of the measures, (3) relate the distribution of the q-FPs to various fibrosis stages and develop a numerical fibrosis scoring system with an expanded dynamic scale of measurement and (4) validate the model in an independent set of biopsies.

MATERIALS AND METHODS

This study was part of a larger human study of the natural history of disease progression and regression performed at Virginia Commonwealth University (VCU IRB 1960). All subjects provided consent to participate in this study. De-identified slides were provided to collaborating investigators at Southern Medical University Nanfang Hospital who performed the dual photon microscopy based q-FP analysis. The q-FP analysis was then related to traditional histopathological analysis performed independently at Virginia Commonwealth University. Investigators at both centers participated in data analysis and interpretation. The authors designed, performed and analyzed the data and take full responsibility for the contents of the manuscript which has been written entirely by the authors.

A: Study population

The study population included subjects with varying phenotypes of NAFLD and controls with normal liver histology as assessed by two established liver pathologists (MJC and MI). Liver tissue sections were taken from tissue blocks obtained at the time of a clinically-indicated liver biopsy. Clinical and laboratory data matching the time of the liver biopsy were extracted from individual subject electronic medical records. Control individuals for this study were identified from those who had a liver biopsy for suspected NAFLD but were found to have <5% steatosis and no other histological lesions including fibrosis on routine evaluation of the biopsy. Subjects with alternate causes of liver disease such as hepatitis C, hepatitis B, hemochromatosis etc. were excluded. All subjects were considered to consume <30 gm of alcohol daily for men and 20 gm of alcohol daily for women based on clinical assessment as done in other studies.^{18, 19}

B: Assessment of routine liver histology

NAFLD was defined by the presence of ≥5% macrovesicular steatosis. Steatohepatitis was defined using NIDDK NASH CRN criteria.⁶ All cases of steatohepatitis had steatosis, inflammation and hepatocellular ballooning. Fibrosis was also staged according to the NASH CRN criteria from stages 0 to 4. For purposes of this analysis, those with definite versus borderline steatohepatitis were considered together and the sub-stages of stage 1 were

all considered as a single stage (stage 1). The rationale for this is that there is yet no evidence that the sub-classifications of stage 1 are linked to clinically meaningful outcomes.

C: Dual photon microscopy measurement of q-FPs

A paraffin-embedded section (4µm thick) of each sample without any specific staining was prepared for dual photon microscopy. Dual photon microscopy imaging and processing by imaging software was performed as previously described.¹⁴ Briefly, the whole section of each slice was imaged with 20 × objective and 512 × 512 pixels of resolution, and then the images of all slices were assessed with computerized image-analysis by two independent investigators (JY and JL) to output the profile of q-FPs data for each slice. The q-FPs that were measured and applied for further analysis are listed in Supplemental Table 1. These included the fibrosis-related parameters such as the collagen content, the geometric and textural features of collagen fibers and the number of cross-linked collagen fibers. The image-analysis software was programmed to measure collagen fibril characteristics in operator-defined segmentation regions of the liver tissue section as described previously (and described in Supplemental Table 1).¹⁴ These included (1) General: the liver section in its entirety, (2) Perisinusoidal: hepatocyte-associated collagen in the perisinusoidal space, (3) Vessel: Collagen fibrils directly connected to veins; and (4) Vessel bridges: Collagen fibrils extending from vein to vein or vein to portal tract. Perisinusoidal fibrosis was defined as the fibrosis signal which is adjacent to hepatocytes in the parenchyma and excluded the signal associated with vessels or septa extending from vessels. The image-analysis software is currently unable to distinguish definitely between central veins and portal veins and vessel-associated collagen was not dependent on the type of vessel. There are thus technical distinctions between q-FP analyses of collagen versus traditional histological methods. Interested readers are referred to prior publications on more technical details of measurement.^{14, 15}

D: Approach to validate q-FP based assessment of fibrosis in NAFLD

Reproducibility of the measurements—Each biopsy was evaluated by two different investigators each on two different days. The concordance data were then computed to determine the intra- and inter-observer agreement for each q-FP. Only q-FPs with an inter-class correlation coefficient (ICC) ≥ 0.8 were used for further analysis.

Defining the spectrum of changes in q-FPs across the fibrosis spectrum of NAFLD—The distribution of q-FP scores across the full spectrum of NAFLD-related fibrosis was measured in the test cohort. The mean value for each q-FP was normalized to a value between 0–1 by relating it to the total score for that q-FP across all stages. The normalized mean scores were visualized in a heat map to compare the general distribution of changes across different stages of disease and different regions (general, perisinusoidal, vessel and vessel-bridges) of the liver sections. The change in q-FP scores between adjacent stages, as defined by $(q\text{-FP}_{n+1} - q\text{-FP}_n) / q\text{-FP}_n$, was used to further visualize those that changed progressively with increasing disease stage. This was performed for general, vascular, perisinusoidal regions and vascular bridges separately.

Concordance with traditional histological assessment—Spearman's correlation coefficients were determined to relate each q-FP to traditional fibrosis stage in the test cohort. Of those that were significantly related, a subset that were not correlated with each other in terms of their operational definitions were selected to avoid co-linearity and entered into a multivariable ordinal logistic regression analysis to develop a model of q-FP evolution in NASH. The area under receiver-operating-characteristic curve (AUROC) for the overall model was then computed. In an alternate analysis, a model was developed from a principal component analysis (PCA) of q-FPs that were significantly correlated with fibrosis stage on univariable analysis. The AUROCs were also tested in an independent set of biopsies (validation cohort) and related to the NASH CRN histological stage to confirm the robustness of the model.

Extending the dynamic scale of measurement by a linearized scale of assessment—Validation of q-FP scores against traditional fibrosis stages does not obviate the limitations of traditional staging, especially with respect to an expanded dynamic range with a continuous scale. This was accomplished by application of the Desirability functions using methodology similar to that described by us before.²⁰

Desirability functions can reset and transform the range of scores so that the lower end reflects a biologically desirable state and is given a score of 0, while the upper end reflects a very undesirable state and given a score of 1. Thus for NASH, a score of 0 would be reflective of absence of fibrosis while a score of 1 would reflect the presence of cirrhosis. The best discriminant score separating stages 0 and 1 based on the regression analysis described above was set as 0. Conversely, the score discriminating stage 3 vs 4 was set as 1. All individual scores below the threshold for 0 were scored as 0, while all scores above the threshold for 1 were scored as 1.

Scores for 4 q-FP parameters found to be significantly related to conventional fibrosis stages from the multivariable regression model described above were converted to a desirability function between 0–1. These individual Desirability functions were then combined using the geometric mean to arrive at a composite measure of the overall index, such as

$$\text{q-FP Desirability Index} = (d_1 \times d_2 \times \dots \times d_4)^{1/4}$$

This thus provided a composite desirability index with an infinite scale for measurement of dynamic changes in fibrosis. Interested readers are also referred to detailed reviews of this methodology for the mathematical background and validity of this approach.^{20, 21}

E: Sample size and statistical analysis

Sample size assessment was done with NQuery (version 4.0). Based on a correlation coefficient of 0.8 or higher and setting the p value at 0.05 with two-sided testing along with a ratio of samples in the gold standard versus new method of 1, a sample size of 40 would provide about 100% power. Since we were unsure of the correlation when we started, a sample size of 50 was used for the model building cohort.

Standard descriptive statistics were used to describe the populations and quantify q-FPs. The choice of statistical tests used to compare groups was based on the type of measure and its distribution. Intra- and inter-observer agreement was evaluated by inter-class concordance (ICC) analysis. ICC was also used for identifying the concordance between q-FP measures of the virtually decreased sample length (i.e., 10 mm and 5 mm, respectively) versus the original length of biopsies. Spearman rank correlation coefficient was used for correlations between variables. Multivariable analysis following univariable logistic regression was used to test the q-FPs independently associated with categorical variables. PCA was performed and provided the variables for the PCA-based model. Two-sided p values <0.05 were considered significant.

RESULTS

Liver biopsy slides from 50 subjects were used as the test cohort. The demographic, clinical and general histological findings from these subjects are shown in Table 1. The subjects were largely overweight or obese and insulin resistant. Individuals with a fatty liver did not have any fibrosis (stage 0). Of those with steatohepatitis, the fibrosis stages were more or less equally distributed from stage 0 to stage 4 (cirrhosis). Thus the entire spectrum of fibrosis associated with NASH was present in the cohort. A second set of 42 biopsies from an independent group of subjects were used to validate the model. They broadly resembled the test cohort with respect to demographic, clinical and histological findings.

1. Reproducibility of the measurements and observer variability

Q-FP parameters noted in Supplemental Table 1 were measured as described previously.¹⁴ To evaluate the test-retest variability in these measures, the intra- and inter-observer concordances for the measures were analyzed from ten randomly selected liver biopsy samples encompassing all stages of fibrosis. In ICC analysis, average inter- and intra-observer concordances were 0.87 ± 0.30 and 0.87 ± 0.29 , respectively. We then set a criterion of ICC = 0.80 for selecting the q-FP to be used for further model development. A total of 70 q-FPs were qualified, with the final average inter- and intra-observer agreement of 0.97 ± 0.03 and 0.98 ± 0.03 , respectively (Supplemental Table 2).

To further evaluate the impact of biopsy length on the assessment of q-FPs, virtual sections of individual biopsies with decreasing size were assessed in comparison with the ones of their original size. Data for the 70 q-FPs remained concordant (ICC 0.95 ± 0.03) even when the length of virtual sections decreased to 10 mm from 21 mm (12 mm–72 mm). When the length decreased to 5 mm, the concordance dropped to 0.83 ± 0.22 .

2. The spectrum of q-FPs across the range of fibrosis in NAFLD

The overall amount and pattern of collagen visualized by dual photon microscopy was similar to traditional histological staging of fibrosis (Figure 1). The normalized mean values for the 70 q-FPs selected are shown in Figure 2. There was overlap in parameters between stages 1 and 2 probably reflecting the presence of perisinusoidal fibrosis in both stages 1 (perisinusoidal or periportal fibrosis) and 2 (periportal and perisinusoidal fibrosis).⁶ Visually, there was a distinct separation with advancing fibrosis stages. The q-FP measures were not

significantly different from stage 1a to 1b and 1c although there tended to be higher q-FP values associated with vessels in stage 1c (Supplemental Figure 1). At the other end of the fibrosis spectrum, some differences in the zonation-based subtypes of stage 4 fibrosis were evident with higher collagen fibril content, number, length, thickness and cross-link numbers at the vessel region in zone 1 versus zone 3 dominant cirrhosis ($p=0.02-0.04$) (Supplemental Figure 2).

3. Specific changes in q-FPs with stage transitions

Next, the changes in q-FPs, defined as $(q-FP_{n+1}-q-FP_n)/q-FP_n$, between the adjacent fibrosis stages were evaluated (Figure 3). From stages 0 to 1, there was an increase in the levels of q-FPs in the general and vessel regions; three (4.3%) q-FPs had a significant increase ($p<0.05$ for all), mainly including the q-FPs of collagen fibril number in vessel region. From stages 1 to 2, the increase was still mainly in the q-FPs in general section and vessel region; three (4.3%) of total q-FPs mainly related to perisinusoidal collagen fibrils were also significantly increased ($p<0.05$ for all). While some septal parameters changed, this did not reach significance. A total of 44 (62.8%) q-FPs changed significantly ($p<0.05$ for all) with progression to stage 3; these included all geometric texture features of collagen fibrils in the vessel-bridges and sinusoidal regions and the number and texture features in the general section. Progression to cirrhosis was associated with further changes in 25 (35.7%) remaining q-FPs ($p<0.05$ for all). Overall, 47 (67.1%) q-FPs had a 1.8-fold (1.1–3.2) increase ($p<0.01$ for all) from stages 1 to 4.

4. Concordance with traditional histological assessment

Fifty-three (75.7%) of the 70 q-FPs analyzed were significantly correlated with conventional fibrosis stages ($p<0.01$ for all, $\gamma_s=0.75-0.35$) (Supplemental Table 3). Among these, 16 q-FPs (22.8%) with the strongest correlation to fibrosis stage (γ_s 0.70, $p<0.001$ for all) were selected for the univariable logistic regression analysis (Supplemental Table 4). Notably, q-FP of collagen content did not reach the correlation standard (γ_s 0.70) and thus was not included in the 16 q-FPs, indicating its inadequacy for reflecting the full range of fibrosis progression as documented.²²

In the first model, a PCA index was developed using the 16 q-FPs with strongest correlation to fibrosis stage. Regression analysis of PCA index indicated its ability to significantly predict fibrosis stages (Odds ratio [OR] 7.8, 95% confidence interval 3.4–18.0, $p<0.001$). It was also able to separate any stage of fibrosis from no fibrosis and cirrhosis from lower stages with AUROCs of 0.88 ($p=0.01$) and 0.93 ($p<0.001$), respectively (Table 2).

Next, a multivariable ordinal logistic regression was performed using parameters significant on univariable analysis that were unrelated in terms of the physical property being measured. These identified four q-FPs including, the number of strands (OR 8.5, $p=0.004$), strand eccentricity (OR 8.3, $p=0.004$), strand solidity (OR 8.0, $p=0.003$) in the general section and the strand length (OR 12.0, $p=0.02$) in vessel region as the parameters that were independently related to fibrosis stage (Supplemental Table 5). The distributions of these parameters with increasing stage are shown in Figure 4 and Supplemental Table 6.

The AUROCs ranged from 0.81 to 0.92 and were all significant for evaluating the ability of the four q-FPs to correctly identify various stages of fibrosis (Table 2). The AUROCs for the minimal model were not significantly different from those of the PCA model described above.

5. Extending the dynamic scale of measurement

Then, the four q-FPs with significance on the multivariable regression analysis were used to develop Desirability functions where a score of 0 reflected the mean scores with a traditional fibrosis stage 0, i.e. highly desirable, whereas the mean score for stage 4 was given a score of 1, i.e. highly undesirable. The desirability scores as a function of parameter score are shown in Figures 5A–D. The overall q-FP Desirability index demonstrated a linear relationship to traditional fibrosis stage (Figure 5E).

6. Assessment of the multivariable model and Desirability functions in validation cohort

In an independent set of liver biopsies (n=42) scored by an independent pathologist with no access to q-FP data, the q-FP scores were computed by the investigators. The ability of the four q-FPs to discriminate between fibrosis stages 0–4 was reconfirmed as shown in Figure 6; the corresponding AUROCs are summarized in Supplemental Table 7. The q-FP Desirability index was also linearly and strongly correlated ($p < 0.0001$) with the traditional fibrosis stages in this set of biopsies (Figure 5E). Furthermore, the slopes of the regression lines for the test and validation cohorts were not significantly different (pooled slope = 0.17/stage).

An additional set of four subjects who had two or three serial biopsies over a period of 4–8 years were also evaluated to probe the ability to tell the differences in fibrosis between biopsy pairs. The pairs were staged to be 3 vs 3, 1 vs 1 vs 1, 2 vs 3, and 3 vs 3, respectively. The four q-FPs all increased to a greater degree in the subject with an increase in fibrosis stage as compared with those who maintained their fibrosis stage; meanwhile, they also detected fine differences in fibrosis even within the same stage for the individual pairs, suggesting that the q-FPs might be sensitive to read the serial changes (Supplemental Figure 3).

DISCUSSION

The assessment of hepatic fibrosis, as a marker of disease progression or regression, has assumed center-stage in the evaluation of NASH.²³ The current study represents an effort to use dual photon based assessment of hepatic fibrosis as a further refinement of traditional histological assessment. The ability to perform partially-computerized assessment of q-FPs also carries the potential to bring some degree of automation to the process and thereby facilitate the overall assessment of liver histology by the pathologist. It is however important to note that histological assessment of NASH involves evaluation of all the other features of NASH including disease activity which require assessment by a pathologist.⁷

An advantage of q-FP based assessment of hepatic fibrosis is the robustness of its accuracy over a range of liver biopsy lengths. One of the principal requirements for high quality traditional histological assessment of the liver is the need for an adequate liver biopsy

sample size.²⁴ Indeed many current trials mandate a biopsy length of 2 cm or greater for subjects to be included. However, in reality, the core size obtained is often smaller than this, and many published studies have reported smaller liver biopsy sizes limiting the reliability of the histologic reads.^{18, 19} In this study, multiple smaller sections of the scanned tissue were evaluated and the results were found to be remarkably reproducible. The q-FP method thus overcomes a potential barrier to more accurate assessment of fibrosis when the tissue sample obtained is less than 1.5–2 cm in length.

Another advantage of the q-FP methodology is that it provides a quantitative result along a continuous numerical scale using the desirability functions used in this study. Such functions have been used as quality assessment tools in industry and are now beginning to be evaluated in medical contexts.^{21, 25} The extended dynamic range may allow assessment of more nuanced changes in fibrosis that may not be captured by traditional scales given their limited range of measurement. This may potentially allow faster detection of therapeutic benefit in clinical trials. Furthermore, by measuring fibrosis along a continuous scale overcomes the errors inherent in assessment of fibrosis using ordinal cutoffs.

In an era of intense ongoing research to develop noninvasive assessment of disease status for NASH, some may question the need for refinement of existing histological methods of fibrosis assessment. It is worth noting that none of the noninvasive methods have met evidentiary standards of biomarker qualification allowing them to be approved specifically for assessment of hepatic fibrosis and that their development is tied to traditional histological assessment with its accompanying strengths and weaknesses. The refinement of the conventional histological assessment to overcome its current limitations will not only potentially benefit therapeutic development efforts but also benefit development of biomarkers by being a better “gold standard”.

In the future, based on this initial step of cross-sectional study, the q-FP score will require longitudinal validation against clinical outcomes to define its specific use. Furthermore, it may permit an alternate way to evaluate the range of fibrosis evolution especially from bridging fibrosis to cirrhosis. If short-term changes along an extended range of fibrosis assessment can be linked to reduced risk of cirrhosis development, such changes could be used as a surrogate endpoint to establish treatment benefit and shorten the time needed for drug development. This is of course hypothetical and a direction for future studies. Similarly, the findings of collagen directly related to vascular and biliary structures corroborate and extend prior findings²⁶ and provide further impetus to the assessment of central and portal veins as well as biliary epithelium in disease progression in NASH.

In summary, q-FP can be used as an objective, reliable measure of fibrosis stage along a numerical scale in subjects with NASH. While more work needs to be done to validate and refine the model using much larger data sets, this study provides proof of concept for the feasibility of developing a quantitative numerical histological scale for assessment of fibrosis in NASH using dual photon microscopy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support:

- T32 DK 61731 to Dr. Sanyal
- RO1 AA020758 to Dr. Sanyal
- RO1 DK 10596 to Dr. Sanyal
- NSFC grant (81371603, 81670522) to Dr. Wang
- GZST grant (2013B051000051) to Dr. Wang

List of Abbreviations

NAFLD	nonalcoholic fatty liver disease
NASH	nonalcoholic steatohepatitis
NASH CRN	NASH Clinical Research Network
q-FP	quantitative fibrosis-related parameter
ICC	intra-class correlation coefficient
AUROC	area under receiver-operating-characteristic curve
PCA	principal component analysis
OR	odds ratio
CI	confidence interval
NoStr	strand number
StrEccentricity	strand eccentricity
StrSolidity	strand solidity
StrLengthV	strand length in vascular region
BMI	body mass index
ALT	alanine aminotransferase
AST	aspartate aminotransferase

References

1. Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA. Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. *Gastroenterology*. 2011; 141:1249–1253. [PubMed: 21726509]

2. 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. [PubMed: 22488764]
3. Sanyal AJ, Friedman SL, McCullough AJ, Dimick-Santos L, et al. American Association for the Study of Liver D, United States F. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American Association for the Study of Liver Diseases-U.S. Food and Drug Administration Joint Workshop. *Hepatology*. 2015; 61:1392–1405. [PubMed: 25557690]
4. Sanyal AJ, Neuschwander-Tetri BA, Tonascia J. End points must be clinically meaningful for drug development in nonalcoholic fatty liver disease. *Gastroenterology*. 2016; 150:11–13. [PubMed: 26602217]
5. Younossi ZM, Stepanova M, Rafiq N, Makhlof H, Younoszai Z, Agrawal R, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. *Hepatology*. 2011; 53:1874–1882. [PubMed: 21360720]
6. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005; 41:1313–1321. [PubMed: 15915461]
7. Brunt EM. Nonalcoholic fatty liver disease: Pros and cons of histologic systems of evaluation. *Int J Mol Sci*. 2016; 17:97.
8. Younossi ZM, Gramlich T, Liu YC, Matteoni C, Petrelli M, Goldblum J, et al. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. *Mod Pathol*. 1998; 11:560–565. [PubMed: 9647594]
9. Nobili V, Parkes J, Bottazzo G, Marcellini M, Cross R, Newman D, et al. Performance of ELF serum markers in predicting fibrosis stage in pediatric non-alcoholic fatty liver disease. *Gastroenterology*. 2009; 136:160–167. [PubMed: 18992746]
10. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009; 7:1104–1112. [PubMed: 19523535]
11. Kim KM, Choi WB, Park SH, Yu E, Lee SG, Lim YS, et al. Diagnosis of hepatic steatosis and fibrosis by transient elastography in asymptomatic healthy individuals: a prospective study of living related potential liver donors. *J Gastroenterol*. 2007; 42:382–388. [PubMed: 17530363]
12. Dulai PS, Sirlin CB, Loomba R. MRI and MRE for non-invasive quantitative assessment of hepatic steatosis and fibrosis in NAFLD and NASH: Clinical trials to clinical practice. *J Hepatol*. 2016; 65:1006–1016. [PubMed: 27312947]
13. Decaris ML, Li KW, Emson CL, Gatmaitan M, Liu S, Wang Y, et al. Identifying nonalcoholic fatty liver disease patients with active fibrosis by measuring extracellular matrix remodeling rates in tissue and blood. *Hepatology*. 2016; doi: 10.1002/hep.28860
14. Xu S, Wang Y, Tai DC, Wang S, Cheng CL, Peng Q, et al. qFibrosis: a fully-quantitative innovative method incorporating histological features to facilitate accurate fibrosis scoring in animal model and chronic hepatitis B patients. *J Hepatol*. 2014; 61:260–269. [PubMed: 24583249]
15. Wang TH, Chen TC, Teng X, Liang KH, Yeh CT. Automated biphasic morphological assessment of hepatitis B-related liver fibrosis using second harmonic generation microscopy. *Sci Rep*. 2015; 5:12962. [PubMed: 26260921]
16. D'Ambrosio R, Aghemo A, Rumi MG, Ronchi G, Donato MF, Paradis V, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. *Hepatology*. 2012; 56:532–543. [PubMed: 22271347]
17. Brunt EM. Nonalcoholic fatty liver disease: what the pathologist can tell the clinician. *Dig Dis*. 2012; 30(Suppl 1):61–68. [PubMed: 23075870]
18. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, Vitamin E, or Placebo for Nonalcoholic Steatohepatitis. *N Engl J Med*. 2010; 362:1675–1685. [PubMed: 20427778]
19. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, Lavine JE, Van Natta ML, Abdelmalek MF, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic

steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet*. 2015; 385:956–965. [PubMed: 25468160]

20. Gennings C, Heuman D, Fulton O, Sanyal AJ. Use of desirability functions to evaluate health status in patients with cirrhosis. *J Hepatol*. 2010; 52:665–671. [PubMed: 20346534]
21. Chen HW, Wong WK, Xu H. Data-driven desirability function to measure patients' disease progression in a longitudinal study. *J Appl Stat*. 2016; 43:783–795. [PubMed: 26997738]
22. O'Brien MJ, Keating NM, Elderiny S, Cerda S, Keaveny AP, Afdhal NH, et al. An assessment of digital image analysis to measure fibrosis in liver biopsy specimens of patients with chronic hepatitis C. *Am J Clin Pathol*. 2000; 114:712–718. [PubMed: 11068544]
23. Sanyal AJ, Miller V. Regulatory science and drug approval for alcoholic and nonalcoholic steatohepatitis. *Gastroenterology*. 2016; 150:1723–1727. [PubMed: 26924092]
24. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003; 38:1449–1457. [PubMed: 14647056]
25. Rafati H, Mirzajani F. Experimental design and desirability function approach for development of novel anticancer nanocarrier delivery systems. *Pharmazie*. 2011; 66:31–36. [PubMed: 21391432]
26. Richardson MM, Jonsson JR, Powell EE, Brunt EM, Neuschwander-Tetri BA, Bhathal PS, et al. Progressive fibrosis in nonalcoholic steatohepatitis: association with altered regeneration and a ductular reaction. *Gastroenterology*. 2007; 133:80–90. [PubMed: 17631134]
27. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2015; 149:389–397. [PubMed: 25935633]

Appendix

Conflicts of Interest:

Yan Wang	None to report
Robert Vincent	None to report
Jinlian Yang	None to report
Amon Asgharpour	None to report
Xieer Liang	None to report
Melissa J. Contos	None to report
Michael O. Idowu	None to report
Mohammed S. Siddiqui	None to report
Faridoddin Mirshahi	None to report
Kalyani Daitya	None to report
Arun J. Sanyal	None for this project. Dr. Sanyal is President of Sanyal Biotechnology and has stock options in Genfit. He has served as a consultant to AbVie, Astra Zeneca, Nitto Denko, Nimbus, Salix, Tobira, Takeda, Fibrogen, Lilly, Zafgen, Novartis, Pfizer, Immuron, Exhalenz and Genfit. He has been an unpaid consultant to Intercept, Echosens, Immuron, Amarin, Ardelyx, Fractyl, Syntlogic, Nordic Bioscience and Bristol Myers Squibb. His institution has received grant support from Gilead, Salix, Tobira, Intercept, Merck, Astra Zeneca and Novartis.

Roles in project:

Name	Design	Performance	Analysis	Write up
Yan Wang	+	+	+	+
Robert Vincent	+	+	+	+
Jinlian Yang		+	+	
Amon Asgharpour	+	+		

Name	Design	Performance	Analysis	Write up
Xieer Liang		+		
Michael O. Idowu		+		
Melissa J. Contos		+		
Kalyani Daitya		+		
Mohammed S. Siddiqui		+		
Faridoddin Mirshahi		+		
Arun J. Sanyal	+	+	+	+

Introductory Statement

Dual photon microscopy based measurement of collagen fibril-related parameters provides a novel method to quantify hepatic fibrosis along a continuous scale with high fidelity. The current study represents a cross-sectional validation of this method against the NIDDK NASH CRN staging system for nonalcoholic fatty liver disease.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1. Images by dual photon microscopy of clinical samples at NASH CRN fibrosis stages F0-4

Second harmonic generation/two-photon excitation fluorescence imaging by dual photon microscope performed a sensitive and specific detection of fibrous collagen disposition (green) in hepatic parenchyma (red) (20×objective, 516×516 pixels resolution/tile). F0 shows minor amounts of perisinusoidal collagen and normal vessel-associated collagen. F1 shows deposition of collagen fibers along the sinusoids. F2 shows incremental deposition of fibrous collagen in both perisinusoidal and periportal areas. F3 shows collagen fibers extending to bridge the vessels. F4 shows the distortion of parenchyma with nodules formed by septa bundles extending from perisinusoidal and vessel-associated collagen fibers.

Figure 2. Heat map of mean level of quantitative fibrosis-related parameters (q-FPs) across fibrosis stages 0–4 in NASH biopsies

q-FPs in general, vessel, vessel-bridges and perisinusoidal regions were calculated into mean values ranging from 0 to 1 normalized by relating the level of their original values to the values of the particular q-FP at all the stages (F0-4). The quantitative profile of q-FPs shows an incremental trend of fibrosis content and the fibrotic pattern of progressing NASH. Data for the test cohort are shown.

Figure 3. Change in levels of quantitative fibrosis-related parameters (q-FPs) between adjacent fibrosis stages in NASH biopsies
Panel A illustrates the change in each q-FP, as defined by $(q-FP_{n+1}-q-FP_n)/q-FP_n$, between adjacent fibrosis stages as a heat map. Column F0 vs F1 shows an increase in fibrosis level of general lobule and perivessel region had the principal change with increase in fibrosis stage; column F1 vs F2 shows the similar trend as that of F0 vs F1 with additional increasing fibrosis in vessel-bridging and perisinusoidal regions; column F2 vs F3 demonstrates that perisinusoidal and vessel-bridging regions had the greatest increase with progression to bridging fibrosis; column F3 vs F4 shows further incremental fibrosis in perivessel and vessel-bridge regions. Panel B shows the number of q-FPs with statistically significant change between the adjacent stages. The ones with the most significant increase occurred from F2 to F3 (62.9%) and from F3 to F4 (35.7%). Data for the test cohort are shown.

Figure 4. Box-Whisker plots of scores of NoStr, StrEccentricity, StrSolidity and StrLengthV for fibrosis stages 0–4 in NASH biopsies

Data for the test cohort are shown. Box, quartile range; whiskers, extreme range; *, outlier; NoStr, number of strands; StrEccentricity, strand eccentricity; StrSolidity, strand solidity; StrLengthV, strand length in vessel region.

Figure 5. Desirability functions of NoStr, StrEccentricity, StrSolidity, StrLengthV and the combination of them

The mean values of the four fibrosis-related parameters (q-FPs) found to be significant based on logistic regression (described in results section) were plotted against Desirability function for the test cohort. The Desirability function was set at 0 for stage 0 fibrosis indicating a highly desirable state of fibrosis, while the Desirability function for score at threshold of stages 3–4 was given a score of 1 indicating established cirrhosis and thus an undesirable state. The mean q-FP scores for stages 1–3 were given varying desirability scores based on the relative change in mortality risk with incremental stage based on mortality risk per stage.²⁷ The shape of the resultant plot was used to determine the relationship of the q-FP score to Desirability function for the four parameters (panels A–D). Panel E represents the overall q-FP Desirability index (the geometric mean of the individual scores for individual subjects) and its relationship to traditional fibrosis staging. There was a tight linear relationship between the overall q-FP index and fibrosis stage for both the test and validation cohorts. The slopes of the data for the two cohorts were not significantly different. NoStr, strand number; StrEccentricity, strand eccentricity; StrSolidity, strand solidity; StrLengthV, strand length in vessel region.

Figure 6. Receiver operation characteristic curves (AUROC) for NoStr, StrEccentricity, StrSolidity and StrLengthV for discriminating different fibrosis stages for the validation set of NASH biopsies

The AUROCs for NoStr, StrEccentricity, StrSolidity and StrLengthV respectively are: A (F0 vs F1-4): 0.81, 0.81, 0.85 and 0.83; B (F0,1 vs F2-4): 0.83, 0.83, 0.85 and 0.85; C (F0-2 vs F3,4): 0.96, 0.96, 0.96 and 0.96; D (F0-3 vs F4): 0.94, 0.94, 0.98 and 0.91. Supplemental Table 7 has details for the above data. NoStr, strand number; StrEccentricity, strand eccentricity; StrSolidity, strand solidity; StrLengthV, strand length in vessel region.

Table 1

Clinical, laboratory, and histologic data of the patient populations

Variable	Test (n=50)	Validation (n=42)
	n (%) or median (range)	
Age (year)	56 (34–75)	57 (36–80)
Gender (male)	10 (20)	7 (17)
BMI (kg/m ²)	33.8 (22.7–43.0)	34.4 (22.2–45.1)
BMI >30 kg/m ²	40 (80)	34 (81)
Diabetes	21 (42)	20 (48)
Hyperlipidemia	32 (64)	30 (71)
AST (U/L)	49 (14–201)	70 (12–164)
ALT (U/L)	54 (13–255)	84 (24–248)
AST/ALT	0.9 (0.5–4.3)	1.0 (0.4–3.7)
AST/ALT >1	14 (28)	10 (24)
Biopsy length (mm)	16.5 (8.0–72.0)	15.5 (6.0–52.0)
Steatosis		
1	25 (50)	17 (40)
2	13 (26)	13 (31)
3	12 (24)	12 (29)
Inflammation		
0	2 (4)	2 (5)
1	26 (52)	32 (76)
2	22 (44)	8 (19)
3	0 (0)	0 (0)
Ballooning		
0	15 (30)	7 (17)
1	26 (52)	23 (55)
2	9 (18)	12 (29)
Fibrosis stage		
0	9 (18)	5 (12)
1	12 (24)	13 (31)
2	12 (24)	10 (24)
3	7 (14)	12 (29)
4	10 (20)	2 (5)

Note: BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

AUROC analysis of the performance of b-FPs and PCA index in discriminating fibrosis stages 0–4 in the test set of NASH biopsies

Table 2

Discrimination of fibrosis stage	Q-FP	AUROC	95%CI	Sensitivity	Specificity
F0 vs F1-4	NoStr	0.87	0.75–0.99	0.84	0.78
	StrEccentricity	0.87	0.75–0.99	0.84	0.78
	StrSolidity	0.87	0.75–0.99	0.84	0.78
	StrLengthV	0.88	0.75–1.00	0.86	0.78
F0,1 vs F2-4	PCA index	0.88	0.77–0.99	0.88	0.78
	NoStr	0.81	0.68–0.93	0.81	0.71
	StrEccentricity	0.81	0.68–0.93	0.81	0.71
	StrSolidity	0.81	0.69–0.93	0.81	0.71
F0,2 vs F3,4	StrLengthV	0.83	0.71–0.95	0.68	0.86
	PCA index	0.83	0.71–0.94	0.81	0.71
	NoStr	0.89	0.79–0.98	0.90	0.82
	StrEccentricity	0.89	0.79–0.98	0.90	0.82
F0,3 vs F4	StrSolidity	0.89	0.79–0.98	0.90	0.82
	StrLengthV	0.86	0.76–0.97	0.84	0.76
	PCA index	0.89	0.79–0.98	0.80	0.88
	NoStr	0.92	0.84–1.00	0.90	0.81
	StrEccentricity	0.92	0.84–1.00	0.90	0.81
	StrSolidity	0.92	0.84–1.00	0.90	0.81
	StrLengthV	0.91	0.82–1.00	0.90	0.83
	PCA index	0.93	0.85–1.00	0.90	0.86

Note: Q-FP, quantitative fibrosis-related parameter; AUROC, area under receiver-operating-characteristic curve; PCA, principal component analysis; CI, confidence interval; NoStr, strand number; StrEccentricity, strand eccentricity; StrSolidity, strand solidity; StrLengthV, strand length in vessel region.