

RESEARCH ARTICLE

Second harmonic generation microscopy provides accurate automated staging of liver fibrosis in patients with non-alcoholic fatty liver disease

Pik Eu Chang^{1,2*}, George Boon Bee Goh^{1,2}, Wei Qiang Leow^{2,3}, Liang Shen⁴, Kiat Hon Lim^{2,3}, Chee Kiat Tan^{1,2}

1 Department of Gastroenterology & Hepatology, Singapore General Hospital, Singapore, **2** Duke-NUS Medical School, Singapore, **3** Department of Anatomical Pathology, Singapore General Hospital, Singapore, **4** Biostatistics Unit, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

* jason.chang@singhealth.com.sg



OPEN ACCESS

Citation: Chang PE, Goh GBB, Leow WQ, Shen L, Lim KH, Tan CK (2018) Second harmonic generation microscopy provides accurate automated staging of liver fibrosis in patients with non-alcoholic fatty liver disease. PLoS ONE 13(6): e0199166. <https://doi.org/10.1371/journal.pone.0199166>

Editor: Manlio Vinciguerra, University College London, UNITED KINGDOM

Received: February 6, 2018

Accepted: June 1, 2018

Published: June 20, 2018

Copyright: © 2018 Chang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research work was supported by a project collaboration agreement with Diagnostics Development (DxD) Hub of Exploit Technologies Pte. Ltd., Singapore, and HistoIndex Pte. Ltd., Singapore. Collaborating parties invested contributions to enable completion of the research. HistoIndex Pte. Ltd. provided the imaging

Abstract

Background

Assessment of severity of liver fibrosis is essential in the management of non-alcoholic fatty liver disease (NAFLD). Second Harmonic Generation (SHG) microscopy is a novel optical tissue imaging system that provides automated quantification of fibrosis based on unique architectural features of collagen. This study aims to develop and validate a SHG-based index for automated staging of liver fibrosis in patients with NAFLD.

Methods

SHG microscopy was performed on archived liver biopsy specimens from 83 patients with NAFLD. A unique algorithm was developed to identify specific SHG parameters that correlated with fibrosis stage. The accuracy of the algorithm was compared against clinical assessment by experienced liver histopathologists using the Brunt fibrosis staging and further validated using the leave-one-out cross-validation method.

Results

Mean age of the study cohort was 51.8 ± 11.7 years, with 41% males. A fibrosis index (SHG B-index) was developed comprising 14 unique SHG-based collagen parameters that correlated with severity of NAFLD fibrosis in a continuous fashion. The SHG B-index had excellent correlation with Brunt fibrosis stage (Spearman's correlation 0.820, $p < 0.001$). AUROCs for prediction of Brunt fibrosis stages 1, 2, 3 and 4 were 0.853, 0.967, 0.985 and 0.941 respectively. In the cross-validation analysis, the SHG B-index demonstrated high specificity for diagnosis of all grades of fibrosis. A SHG B-index score of > 1.76 had an overall diagnostic accuracy of 98.5% for prediction of presence of bridging fibrosis (Brunt stage ≥ 3) with sensitivity of 87.5%, specificity 98.0%, positive predictive value 96.6% and negative predictive value 92.6%.

hardware and assisted in the development of the algorithm. Diagnostics Development Hub provided funding to support retrieval of the archived biopsy specimens, collation of clinical data and development and validation of the algorithm.

Employees of HistoIndex Pte. Ltd. assisted in the description of the study methodology, provision of images and assisted in data analysis. The funders otherwise had no role in data collection, decision to publish or preparation of the manuscript.

Competing interests: We have the following interests: This research work was supported by a project collaboration agreement with Diagnostics Development (DxD) Hub of Exploit Technologies Pte. Ltd., Singapore, and HistoIndex Pte. Ltd., Singapore. Collaborating parties invested contributions to enable completion of the research. HistoIndex Pte. Ltd. provided the imaging hardware and assisted in the development of the algorithm. Diagnostics Development Hub provided funding to support retrieval of the archived biopsy specimens, collation of clinical data and development and validation of the algorithm. Employees of HistoIndex Pte. Ltd. assisted in the description of the study methodology, provision of images and assisted in data analysis. There are no patents, products in development or marketed products to declare. This does not alter our adherence to all the PLOS ONE policies on sharing data and materials.

Conclusion

The SHG B-index is a unique SHG-based index that provides accurate automated assessment of fibrosis stage in NAFLD patients.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a spectrum of disorders characterized by excessive accumulation of hepatic fat associated with insulin resistance [1]. The prevalence of this disease is increasing worldwide in tandem with the obesity epidemic in developed countries. It is currently the most common liver condition in Western countries, affecting 17–46% of the adult population [2]. Worryingly, the prevalence of NAFLD exceeds 90% in patients with severe obesity and up to 69% in patients with type 2 diabetes mellitus [3]. NAFLD is histologically characterized into nonalcoholic fatty liver (NAFL) and nonalcoholic steatohepatitis (NASH) [4]. While NAFL is generally benign with low risk of progression of liver disease, NASH covers a wide spectrum of disease activity including fibrosis, cirrhosis and hepatocellular carcinoma (HCC) [1,2]. The severity of fibrosis is the most important prognostic factor in NAFLD and is the strongest predictor for the development of liver cirrhosis, need for liver transplantation, development of HCC and liver-related death [5,6]. Assessment of liver fibrosis is thus essential in the management of NAFLD.

The current gold standard for assessing liver fibrosis is with liver biopsy [7]. In the management of NAFLD, liver biopsy is essential both for the diagnosis of NASH and for fibrosis staging [8]. While serum biomarkers and transient elastography are acceptable non-invasive tests for screening of NAFLD patients with low risk of advanced fibrosis, reliable identification of at-risk patients with advanced fibrosis still requires confirmation by liver biopsy [2,9]. Unfortunately, liver biopsy is an imperfect gold standard due to sampling variability and inter-observer variability [10–13]. The NASH Clinical Research Network (NASH-CRN) system is the current reference histological scoring system used for assessment of fibrosis in NAFLD [14]. Severity of fibrosis is based on the distribution of fibrosis with initial peri-sinusoidal fibrosis (stage 1), progressing to portal/peri-portal fibrosis (stage 2), bridging fibrosis (stage 3) and finally cirrhosis (stage 4) [15]. Despite widespread adoption of this histological scoring system, poor inter-observer agreement remains a challenge [16]. Furthermore, the semi-quantitative nature of the fibrosis scoring system creates potential for error when the differentiation between groups is not distinct. Given that the process of fibrogenesis and fibrolysis is dynamic, an objective quantitative assessment on a continuous scale provides greater accuracy than a semi-quantitative score, especially when assessing changes in fibrosis over time [17,18]. This has important implications in the current era of antifibrotic therapy where linear changes in fibrosis quantification may be more meaningful to demonstrate response to therapy compared to the need to demonstrate stage migration.

There is thus an unmet need for a robust and accurate automated system of fibrosis assessment that will reduce inter-observer variability in the assessment of NASH fibrosis. In addition, this system should provide a continuous scale of fibrosis assessment that objectively evaluates changes in fibrosis severity. Most therapeutic trials for NASH require demonstration of histological improvement in fibrosis as a hard end-point to demonstrate efficacy. Such an automated fibrosis assessment tool would facilitate wide-scale international multi-center therapeutic trials by providing an efficient and reliable system for assessment of liver fibrosis.

Second harmonic generation (SHG) microscopy is a novel tissue imaging system based on non-linear optical microscopy which enables observation of endogenous tissue signals such as Two-Photon Excitation Fluorescence (TPEF) and Second Harmonic Generation (SHG) in unstained tissue samples [18]. TPEF provides visualization of the background liver architecture while the SHG signal provides accurate identification of fibrillar collagen [19,20]. SHG microscopy has been demonstrated to be reliable for fibrosis staging in chronic hepatitis B (CHB), outperforming collagen proportionate area measurement [21]. Computerized algorithms based on SHG pattern recognition and logic modelling allow automated determination of severity of fibrosis with the added advantage of objectivity, reproducibility and fast turn-around time [22]. However, the SHG fibrosis parameters in NAFLD are different from those in CHB due to differences in the distribution and nature of fibrosis. There is thus a need for NAFLD-specific SHG algorithms that accurately predict fibrosis stage based on Brunt rather than METAVIR fibrosis scoring. Proof-of-concept studies have provided encouraging evidence that support the role of SHG microscopy for evaluation of NASH fibrosis [23,24]. However, the reliability of SHG-based automated assessment for NAFLD fibrosis staging is not well established and requires validation.

We have developed an optical microscopy system that can scan and analyze the SHG properties of collagen in unstained liver tissue specimens. In this study, we aim to develop an algorithm based on SHG properties that are unique to NAFLD fibrosis and to validate its reliability to predict severity of fibrosis in patients with NAFLD.

Methods

Patient biopsy samples

Archived formalin-fixed, paraffin-embedded liver biopsy samples from 83 patients with NAFLD were retrieved for analysis. Subjects were identified from a database of patients in the Department of Gastroenterology and Hepatology of the Singapore General Hospital with a clinical diagnosis of NAFLD who underwent liver biopsy for clinical indications between 2005 and 2015. All subjects fulfilled the definition of NAFLD with >5% hepatic steatosis on liver biopsy in the absence of viral hepatitis, significant alcohol consumption (defined as >21 units of alcohol per week in males and >14 units in females), autoimmune and hereditary liver diseases [4]. Liver biopsy was performed for diagnostic and prognostic indications. The liver biopsy samples were processed as per routine clinical practice. The study protocol was approved by the SingHealth Centralized Institutional Review Board (CIRB Reference No: 2015/2527). 5 μ m-thick sections were cut from archived liver biopsy tissue and the anonymized specimens were sent for SHG imaging. The same sections were subsequently stained with hematoxylin and eosin (H&E) stain for blinded assessment by the histopathologists. Special histochemical stains such as trichrome and Victoria blue for assessment of liver fibrosis were available in the immediate preceding sections. Biopsy samples were reported as per clinical protocol by experienced liver histopathologists and severity of liver fibrosis was scored using Brunt's classification [15].

Image analysis

Unstained tissues of the 83 liver biopsy samples were imaged by the Genesis™ system (HistoIndex Pte. Ltd., Singapore). SHG microscopy was used to visualize collagen and two-photon excited fluorescence (TPEF) microscopy was utilized for visualization of the other cell structures. The settings of laser were similar to previous publications [22]. Images were acquired at 20 \times magnification with 512 \times 512 pixels resolution, and each image tile had a dimension of 200 \times 200 μ m. A previous study had shown that SHG/TPEF technology-based quantification

method was less sensitive to sampling error [22]. Thus, 10 five-by-five multi-tile images were randomly acquired for each sample with a final sampling size of 10 mm² instead of scanning the entire biopsy tissue.

Landmark structures such as the portal tract and central vein were marked out on the H&E-stained slide and correlated to various variables on the SHG image (Fig 1). This allowed us to distinguish fibrosis occurring in 3 separate regions: central vein region (CV), portal tract region (PT) and peri-sinusoidal region (PS). An image processing algorithm was used to quantify fibrosis features in these three specific regions. Image analysis was divided into two parts: (1) determining specific regions, (2) quantifying collagen features in each region. The images were processed and analyzed using MATLAB 8.3 (The MathWork, USA).

1. **Determining specific regions:** The SHG signal and TPEF signal in the background of SHG/TPEF images were recognized as noises, which were firstly removed (Fig 1). Collagen was detected from the SHG channel by the Ostu's automatic threshold method [25]. The holes inside biopsy tissues which were caused by vessels, bile ducts, steatosis and unnatural cracks, were detected from the TPEF channel. Based on the features of each detected hole, such as density of hole, width and length of hole, hole solidity and the area of surrounding collagen of hole, the holes caused by vessels and bile ducts were distinguished from others by a decision tree, which was constructed by classification and regression tree (CART) method [26]. Multiple structures were then constructed based on the distances among the holes and the surrounding collagen of these holes. Each structure represented a portal tract region or a central vein region. Finally, these two regions were separated by another CART with features, such as the number of holes, the total area of holes, the area of maximal hole, and the area of collagen. The rest of the tissue area was regarded as perisinusoidal regions. This allowed the identification of specific collagen distribution at the central vein, portal tract and perisinusoidal areas in various stages of fibrosis (Fig 2).
2. **Quantifying collagen features in each region:** In each region, aggregated collagen and distributed collagen were separated based on the cross-linking properties (Fig 3). Aggregated collagen was cross-linked, whereas distributed collagen was not. A total of 28 features including the percentages of different collagen patterns (all collagen, aggregated collagen and distributed collagen) and collagen string features (such as number of strings, numbers of short/long/thick/thin strings, numbers of short/long/thick/thin aggregated/ distributed strings, etc.) were extracted in each region. We used the relative thickness of collagen strings to each other to differentiate between thick and thin collagen strings. Thick strings were defined as collagen strings with a width/length ratio > 0.25. Another 16 similar features were extracted from the entire biopsy tissue. In total, 100 collagen features were extracted and quantified.

Development of SHG algorithm for assessment of fibrosis

To automate the assessment of fibrosis stages of patients with NAFLD, a prediction model was developed based on the quantified collagen features. Firstly, feature selection was performed to reduce the dimensionality of data by selecting only a subset of quantified features. A common method of feature selection, named sequential feature selection was used in this study [27]. In the procedure of sequential feature selection, linear regression model was used whereby the criterion was residual sum of squares and the search algorithm was sequential forward selection. All the 83 subjects were employed to find the most significant collagen features related with Brunt stages. Fourteen features were finally selected, including the percentage of aggregated collagen (Agg), the width of strings (StrWidth), the percentage of central vein collagen

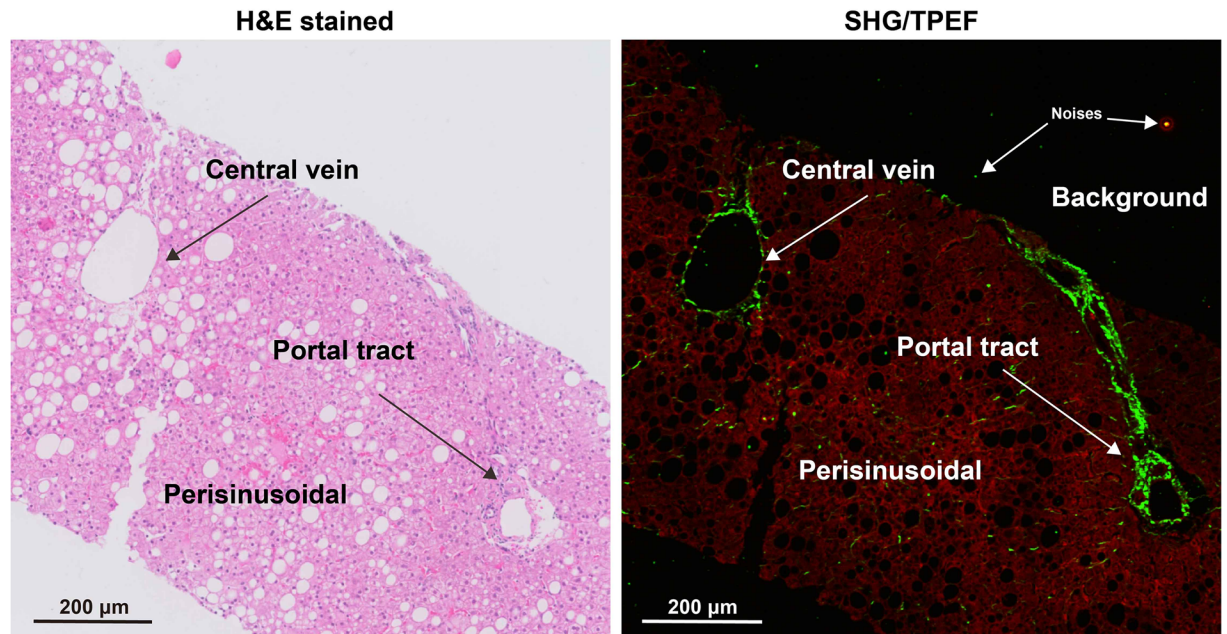


Fig 1. Illustration of central vein, portal tract and perisinusoidal regions in H&E stained image and SHG/TPEF image. H&E: hematoxylin & eosin, SHG: second harmonic generation, TPEF: two-photon excited fluorescence.

<https://doi.org/10.1371/journal.pone.0199166.g001>

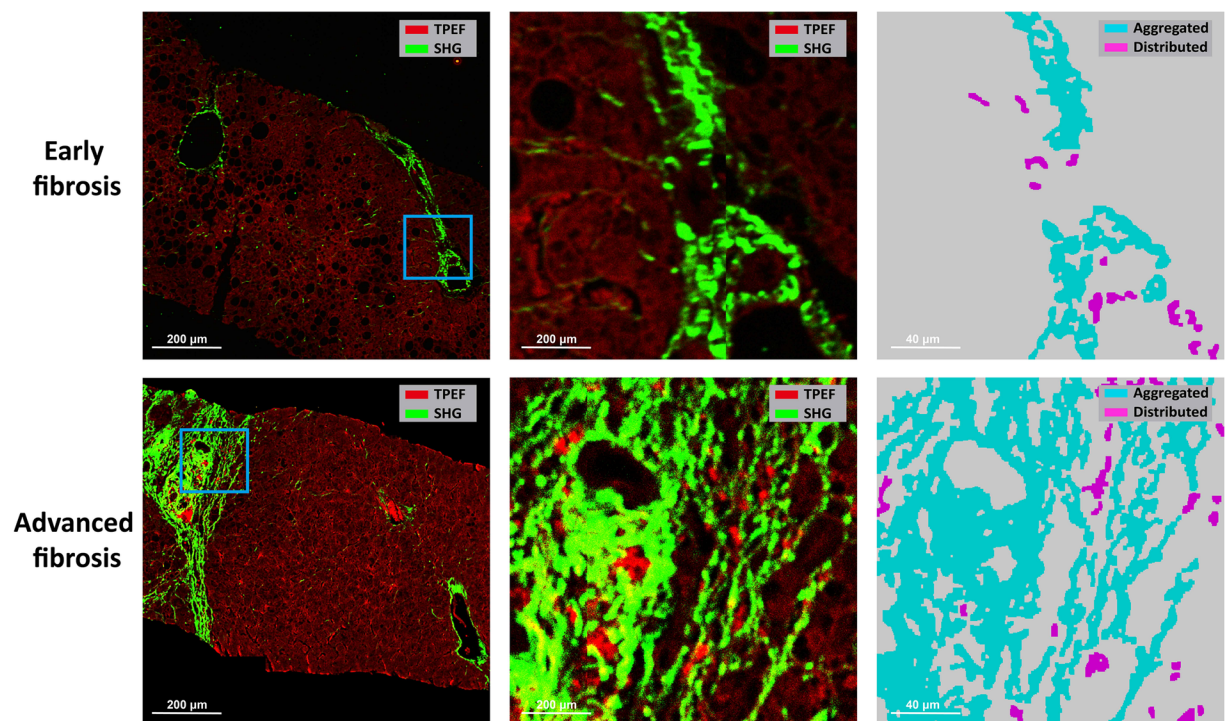


Fig 2. Illustration of specific collagen distribution at central vein, portal tract and perisinusoidal regions in early versus advanced fibrosis.

<https://doi.org/10.1371/journal.pone.0199166.g002>

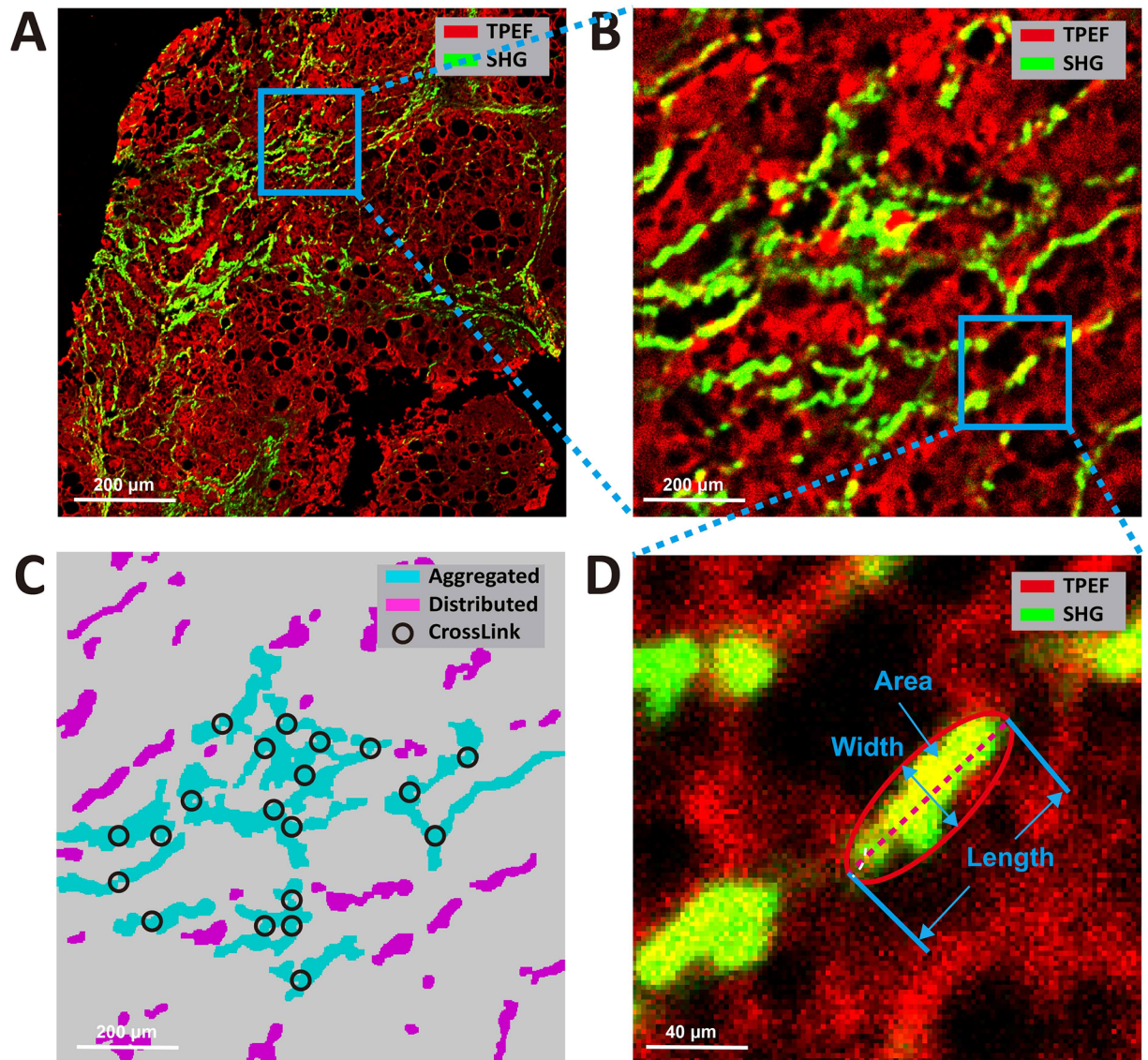


Fig 3. Illustration of collagen parameters on SHG microscopy. A) The RAW SHG/TPEF image. (B) The high magnification imaging performed in the area delimited by the blue square in picture A. (C) The aggregated and distributed strings of picture B. (D) The high magnification imaging performed in the area delimited by the blue square in picture B, and illustration of the width, length and area of a string. The thick string is by definition length-width ratio > 0.25 .

<https://doi.org/10.1371/journal.pone.0199166.g003>

(CV), the number of thick strings in CV (NoThickStrCV), the area of string in CV (StrAreaCV), the length of strings in CV (StrLengthCV), the length of aggregated strings in CV (StrLengthCVA), the area of distributed strings in CV (StrAreaCVD), the number of thin and aggregated strings in PT (NoThinStrPTA), the number of thick and distributed strings in PT (NoThickStrPTD), the number of distributed strings in PS (NoStrPSD), the number of short and distributed strings in PS (NoShortStrPSD), the area of distributed strings in PS (StrAreaPSD) and the length of distributed strings in PS (StrLengthSFD) (S1 Table).

Next, the training process was performed on all the 83 subjects and the validation process was performed using leave-one-out cross-validation method [28]. The selected features were used to assess the severity of liver fibrosis of NAFLD patients using a prediction model called

the SHG B-index, which was constructed from all 83 subjects with multivariable linear regression method. If the result of the prediction model was less than 0, it was set to 0. Thus the severity of liver fibrosis was greater than 0 and may be greater than 4. To estimate how accurately the prediction model will perform in practice, leave-one-out cross-validation method was used (S1 Fig). Briefly, one sample is randomly retained as the validation data while the remaining 82 samples are used as training data to construct the prediction model. The performance of the prediction model is then tested on the single validation sample. The cross-validation process is repeated 83 times, with a different sample left out each time. The overall performance of the prediction model is assessed based on the sensitivity, specificity, positive and negative predictive values on the 83 predicted SHG B-indexes.

Statistical analysis

Clinical and biochemical data were retrieved from the hospital's computer database system. Statistical analysis was done with SPSS version 23 (Chicago, IL, USA). Quantitative fibrosis assessment by the SHG B-index was correlated with Brunt fibrosis score using Spearman correlation. AUROC analysis was performed to evaluate the accuracy of SHG B-index for prediction of the different grades of Brunt fibrosis. Optimal cut-off values of the quantitative SHG B-index were identified for prediction of Brunt fibrosis grades. Leave-one-out cross-validation (described above) was used to validate the performance of the SHG B-index model. The sensitivity, specificity, positive and negative likelihood ratios of SHG B-index for prediction for presence of bridging fibrosis (i.e. Brunt fibrosis grade 3) were calculated. Statistical significance level was set at $p < 0.05$.

Results

Mean age of the study cohort was 51.8 ± 11.7 years, with 41% males and 77% of Chinese ethnicity. Demographic, anthropometric and biochemistry data are summarized in Table 1. Mean biopsy length, calculated by addition of all available fragments, was 2.4 ± 0.9 cm. Distribution of fibrosis in the study cohort was 22.9%, 30.1%, 8.4%, 16.9% and 21.7% for Brunt fibrosis stages B0, B1, B2, B3 and B4 respectively. Analysis of the SHG-derived data identified 14 SHG-based collagen parameters that correlated strongly with the Brunt classification. An algorithm was developed to provide a composite index of SHG-parameters that provided a quantitative assessment of fibrosis severity (S2 Table). This fibrosis index thus derived is

Table 1. Baseline characteristics of NAFLD cohort.

	All subjects (n = 83)
Mean age (years)	51.8 ± 11.7
Male gender	34 (41%)
Chinese ethnicity	64 (77%)
BMI (kg/m^2)	28.4 ± 6.9
HbA1c (%)	6.9 ± 1.2
Albumin (g/L)	42.7 ± 29.0
Bilirubin ($\mu\text{mol}/\text{L}$)	19.6 ± 26.4
Alanine transaminase (U/L)	80.9 ± 56.7
Aspartate transaminase (U/L)	62.3 ± 44.3
Platelet count ($\times 10^9/\text{L}$)	220 ± 83
Prothrombin Time (s)	10.6 ± 1.3
Creatinine (mmol/L)	71.8 ± 24.7

<https://doi.org/10.1371/journal.pone.0199166.t001>

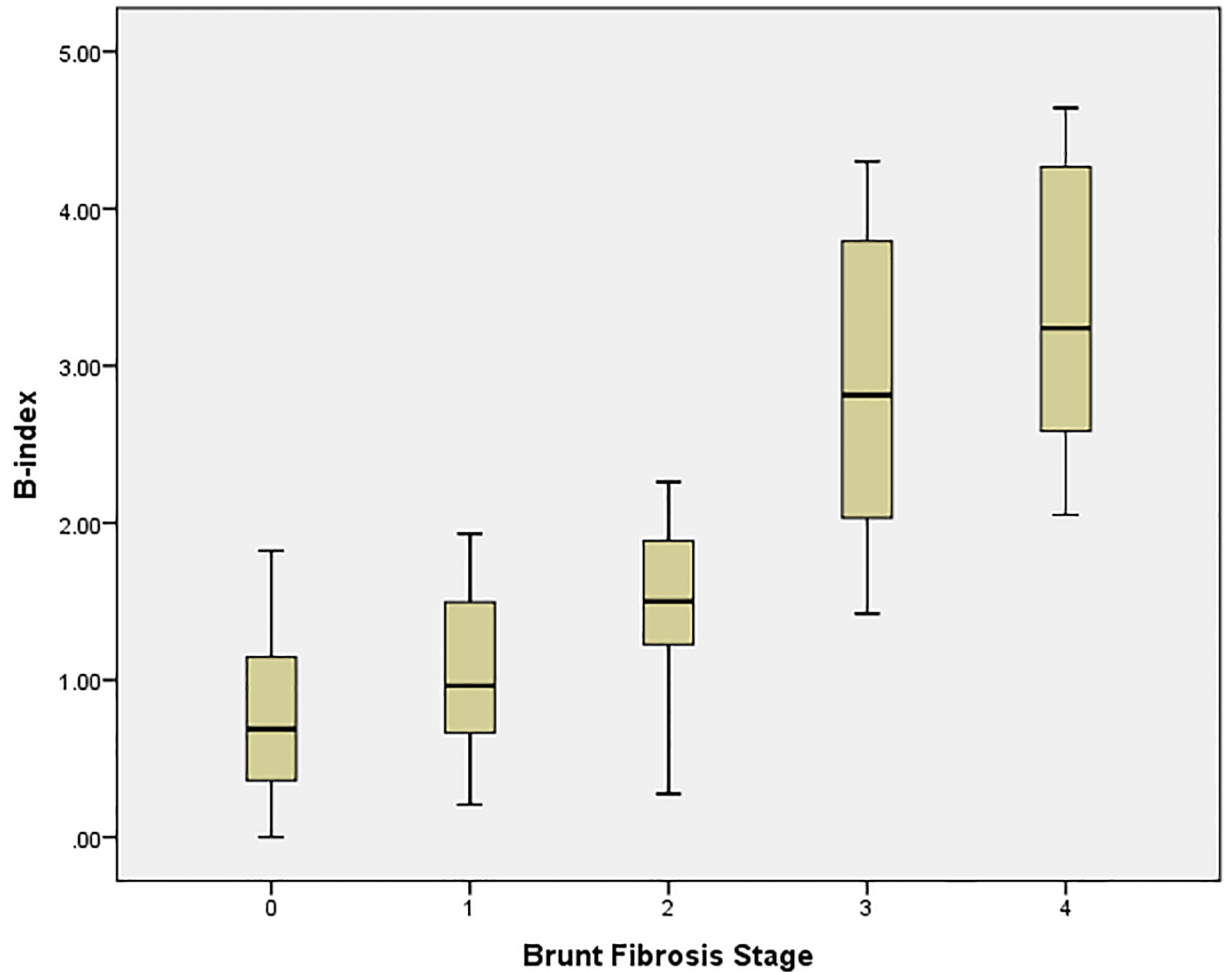


Fig 4. Boxplot of SHG B-index and Brunt fibrosis stage (n = 83).

<https://doi.org/10.1371/journal.pone.0199166.g004>

termed the SHG B-index. The SHG B-index had a high correlation of 0.820 ($p < 0.001$) with Brunt fibrosis staging, with an excellent ability to differentiate advanced fibrosis (Brunt stage 3 and 4) from no or mild fibrosis (Brunt stage 0 and 1). However, there was considerable overlap between Brunt stages 0, 1 and 2 (Fig 4).

AUROC analysis was performed to evaluate the accuracy of the SHG B-index to identify the various fibrosis stages (Table 2). AUROC of SHG B-index to predict mild fibrosis was 0.853 (95% confidence interval 0.774–0.993), significant fibrosis 0.967 (95% confidence interval 0.933–1.000), bridging fibrosis 0.985 (95% confidence interval 0.966–1.000) and cirrhosis 0.941 (95% confidence interval 0.892–0.990) (Fig 5).

Table 2. AUROCs of SHG B-index for prediction of Brunt fibrosis.

NAFLD Fibrosis Group	Brunt fibrosis stage	AUROC	P	95% CI	SHG B-index cut-off value
Mild fibrosis	0 vs 1/2/3/4	0.853	<0.001	0.774–0.933	1.18
Significant fibrosis	0/1 vs 2/3/4	0.967	<0.001	0.933–1.000	1.33
Bridging fibrosis	0/1/2 vs 3/4	0.985	<0.001	0.966–1.000	1.76
Cirrhosis	0/1/2/3 vs 4	0.941	<0.001	0.892–0.990	2.76

<https://doi.org/10.1371/journal.pone.0199166.t002>

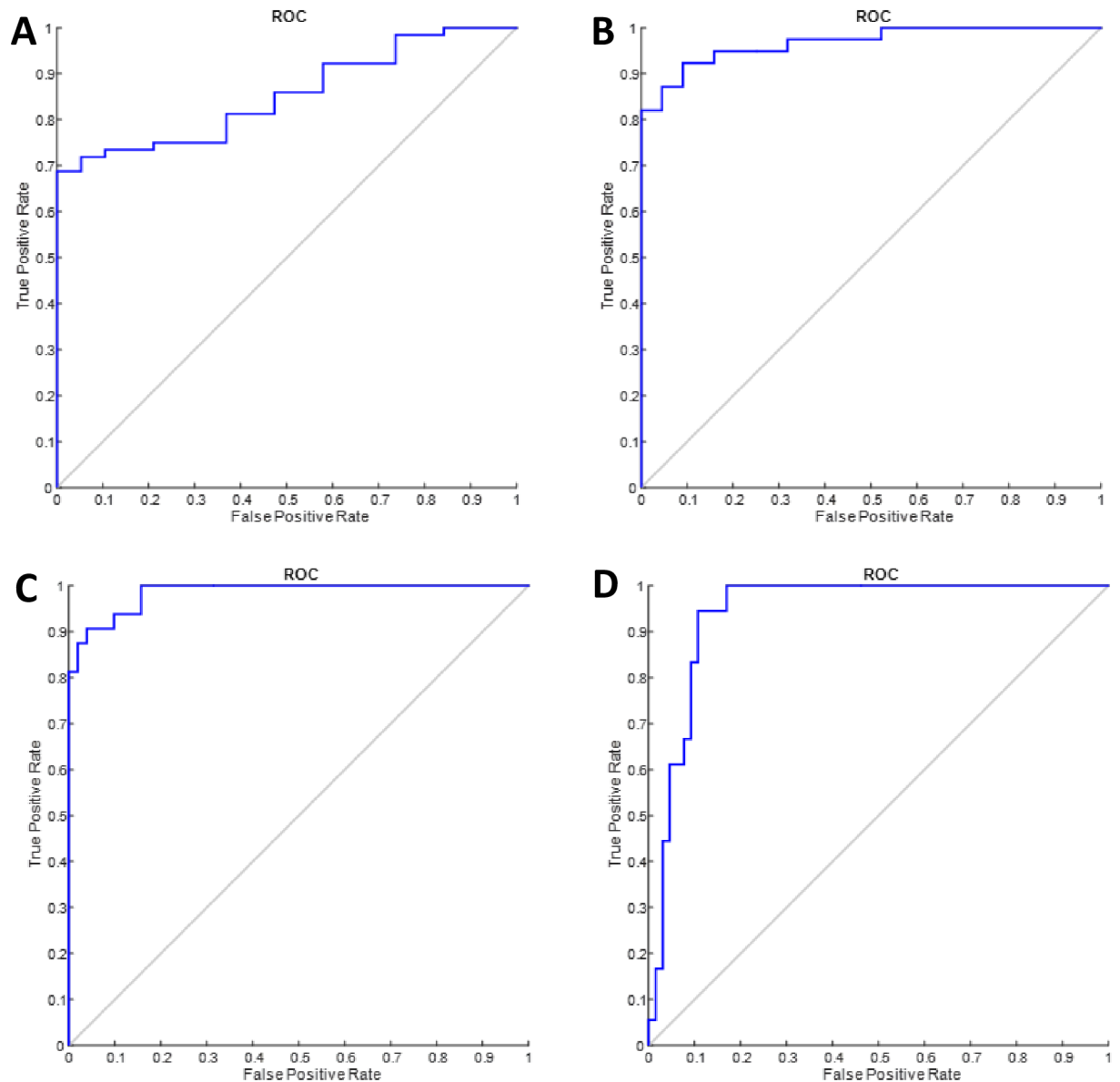


Fig 5. AUROCs of SHG B-index to predict fibrosis stages.

<https://doi.org/10.1371/journal.pone.0199166.g005>

Optimal cut-off values of SHG B-index to identify specific Brunt fibrosis groups were calculated using Youden’s index (maximal sum of sensitivity and specificity). To assess for presence of fibrosis (i.e. to differentiate between no fibrosis and presence of fibrosis), the optimal cut-off value of the SHG B-index was 1.18. The cut-off values for presence of significant fibrosis, advanced fibrosis and cirrhosis were 1.33, 1.76 and 2.76, respectively.

The performance of the SHG B-index was validated using the leave-one-out cross-validation method (Table 3). Overall, the SHG B-index performed very well with high specificity ranging from 84.2% to 98.0% for diagnosis of all grades of fibrosis. The SHG B-index was less sensitive for identification of mild fibrosis but it had a high positive predictive value with a value above 1.18 providing a 93.2% probability of presence of fibrosis. Conversely a SHG B-index value below 2.76 provided a 93.6% probability of not having cirrhosis. The SHG B-index

Table 3. Cross-validation of SHG B-index to predict Brunt fibrosis stage.

NAFLD Fibrosis Group	Brunt fibrosis stage	SHG B-index cut-off	Sens	Spec	+LR	-LR	PPV	NPV
Mild fibrosis	0 vs 1/2/3/4	1.18	65.6%	84.2%	4.15	0.41	93.2%	41.0%
Significant fibrosis	0/1 vs 2/3/4	1.33	84.6%	86.4%	6.22	0.18	86.8%	86.7%
Bridging fibrosis	0/1/2 vs 3/4	1.76	87.5%	98.0%	43.75	0.13	96.6%	92.6%
Cirrhosis	0/1/2/3 vs 4	2.76	83.8%	89.2%	7.71	0.19	66.7%	93.6%

Sens = Sensitivity, Spec = Specificity

+LR = Positive likelihood ratio, -LR = Negative likelihood ratio

PPV = Positive Predictive Value, NPV = Negative Predictive Value

<https://doi.org/10.1371/journal.pone.0199166.t003>

performed best for the prediction of advanced fibrosis, which is particularly relevant in the management of NAFLD. A SHG B-index value greater than 1.76 had an overall diagnostic accuracy of 98.5% for prediction of presence of bridging fibrosis (Brunt stage ≥ 3) with sensitivity of 87.5%, specificity 98.0%, positive likelihood ratio 43.75 and negative likelihood ratio of 0.13, positive predictive value of 96.6% and negative predictive value of 92.6%.

Discussion

Using the unique capabilities of SHG to identify collagen parameters associated with NAFLD fibrosis, we developed a composite index (the SHG B-index) that has high accuracy for identifying various stages of NAFLD fibrosis. Using a cross-validation method, we confirm that the SHG B-index is reliable for the prediction of severity of liver fibrosis compared to traditional histopathologists' assessment using Brunt classification. A SHG B-index of >1.76 is highly sensitive and highly specific for presence of bridging fibrosis, a clinically important feature that is associated with poor prognosis in NAFLD patients. This study thus provides evidence to support the potential clinical application of an automated system of fibrosis staging using unstained liver biopsy specimens.

The development of SHG-based optical microscopy represents a potential paradigm shift in the histopathological assessment of fibrosis, providing several advantages over conventional histopathological assessment. The unique crystalline triple-helix structure of fibrillar collagen provides a specific SHG signal that allows accurate identification, characterization and quantification of liver fibrosis [19]. This process does not require staining of tissues, thus reducing degradation of the sample and production of artifacts which impair the quality of the interpretation [21]. The ability to characterize fibrillar collagen within a 3-dimensional lattice framework provides the opportunity to evaluate the process of fibrosis remodeling, which is a unique advantage over the 2-dimensional assessment using standard histopathology [18]. Once the system has been calibrated, the method is highly reproducible, allowing standardization of measurement and reduction of interobserver variability. SHG-based microscopy thus has the potential to provide an automated assessment of fibrosis stage within hours of placing an unstained liver biopsy slide into a compact machine. From a clinical perspective, it is crucial to develop SHG-based algorithms that provide reliable assessment of fibrosis based on the specific etiology of disease, since current fibrosis staging systems tend to be disease-specific.

SHG-based assessment of liver fibrosis has been shown to be reliable in chronic viral hepatitis. Correlation between SHG signal and METAVIR fibrosis score in patients with chronic viral hepatitis was initially described by Gailhouste and colleagues [21] and subsequently validated in an independent study [22]. However, the distribution of fibrosis in NAFLD is different from that in viral hepatitis. In the former, fibrosis begins in the peri-cellular and peri-sinusoidal areas whereas fibrosis begins in the periportal regions in the latter. The SHG algorithms that

were developed based on METAVIR staging thus cannot be applied to NAFLD. There is thus an important need to develop and validate SHG-based algorithms that are unique for NAFLD, paying particular attention to the ability to geographically differentiate fibrosis deposition in the central vein, portal tract and peri-sinusoidal areas. This is particularly relevant because most clinical liver biopsies are currently performed for NAFLD. As a testament to the developing interest in this area, a proof-of-concept study recently explored the use of SHG microscopy for assessment of fibrosis stage in NAFLD [24]. The authors identified four SHG-based parameters that reliably differentiated the various fibrosis stages with AUROCs of 0.81 to 0.92. However, their algorithm lacked the ability to differentiate portal vein fibrosis from central vein fibrosis. The results of our study thus validate the reliability of SHG as an automated method for fibrosis staging. In addition, our algorithm can localize and quantify central vein fibrosis, portal tract fibrosis and peri-sinusoidal fibrosis. This provides the potential to differentiate fibrosis from different etiologies, for example in patients with concomitant NASH and chronic viral hepatitis. Furthermore, the SHG B-index that we have developed provides a convenient single composite index that can be easily applied in routine clinical practice to provide automated assessment of fibrosis.

We observed that the quantitative distribution of fibrosis assessed by SHG B-index was highly sensitive and specific for the identification of advanced fibrosis in NAFLD but was less discerning in discriminating between early stages of fibrosis. In the early stages of fibrosis (Brunt stage 1 and 2), the amount of collagen in the liver is similar, but the distribution of collagen is different being predominantly peri-sinusoidal in stage 1 and peri-portal in stage 2. However, in advanced fibrosis (Brunt stage 3) we observe a significant increase in the SHG B-index, reflecting an increase in the amount of collagen deposited as the fibrous tissue progresses from the portal tracts to the central veins. The SHG B-index thus provides a reliable method to identify NAFLD patients with advanced fibrosis who are at increased risk of liver-related morbidity and mortality.

SHG microscopy has several important clinical implications in the management of NAFLD patients. Firstly, it provides a highly reliable automated assessment of fibrosis that can reduce interobserver variability and improve accuracy of histological staging of liver fibrosis. Secondly, it can be performed on unstained tissues thus avoiding the need for cumbersome staining procedures and degradation artifacts. Thirdly, SHG optical microscopy provides a quantitative assessment of fibrosis on a continuous scale which reflects dynamic change in fibrosis more accurately than the conventional semi-quantitative histological grading systems. With the increase in new anti-fibrotic therapeutic compounds, the use of a quantitative method with a continuous scoring scale to assess progression or regression in severity of fibrosis is intuitively more reflective of the true therapeutic response. The primary utility of SHG-based assessment of liver fibrosis may be in therapeutic NAFLD trials, many of which require serial liver biopsies for assessment of improvement in fibrosis. The automated fibrosis scoring provided by SHG microscopy helps to minimize inter-observer variability and reduces the cost and logistics associated with shipping of samples for central histology review. This automated, standardized and clinically reliable fibrosis assessment system may thus pave the way for future large, multicenter, international collaborative studies.

Although SHG optical microscopy offers several novel applications in fibrosis assessment, it is important to emphasize that this technology does not replace the need for invasive liver biopsy and cannot replace the key role of the histopathologist in clinical practice [18]. The information provided by the SHG B-index is only limited to fibrosis assessment whereas the histopathologists provides essential details on other features such as inflammation, steatosis, ballooning degeneration, Mallory-Denk bodies, iron deposition and exclusion of other liver diseases. It also does not overcome the problem of sampling error which is an inherent

limitation of liver biopsy. For this reason, we feel that SHG-based fibrosis assessment is most suitable as an adjunct tool that provides additional value to the efficiency and reliability of the histopathologist in clinical practice. The SHG B-index can be used to provide an external reference to guide the histopathologist, thus minimizing inter-observer variability in assessment of fibrosis in NAFLD. This can help to reduce the clinical variability in fibrosis assessment using the NASH-CRN criteria among community-based histopathologists, thus improving the accuracy of fibrosis assessment in general clinical practice [16]. It can also be a useful training tool for less-experienced histopathologists by providing a guide for accuracy and standardization of fibrosis scoring.

We acknowledge that our study is limited by a modest sample size from a single center and lack independent external validation. Nonetheless, we have employed a robust methodology and analysis to provide reliable results, which we believe are important to the scientific community. A larger multi-center is currently being planned to validate the SHG B-index in an independent cohort. Another limitation is that this study did not examine changes in fibrosis stage from subjects with repeat liver biopsies. Hence, the role of SHG B-index for assessment of progression or regression of NAFLD fibrosis in longitudinal pre-and post-treatment liver biopsies needs to be validated as well, particularly in the context of response to pharmacological intervention for NAFLD.

In conclusion, this study demonstrates that the SHG B-index, a unique fibrosis index based on SHG optical microscopy, provides highly reliable automated assessment of fibrosis in patients with NAFLD. This index has a high specificity for advanced fibrosis which is a clinically important sentinel point in the management of NAFLD.

Supporting information

S1 Table. List of SHG parameters.

(DOCX)

S2 Table. Descriptive statistics of B-index prediction model.

(DOCX)

S1 Fig. Supplementary figure: Illustration of leave-one-out cross-validation. One sample is randomly retained as the validation set while the remaining 82 samples are used as the training set to develop a prediction model. The performance of the prediction model is then tested on the single validation sample. The cross-validation process is repeated 83 times, with a different sample left out each time. The overall performance of the prediction model to predict the various fibrosis stages is assessed based on the 83 predicted B-indices.

(PDF)

Acknowledgments

All authors contributed to the planning and execution of the study, data analysis and preparation of the manuscript. The authors would like to acknowledge the contributions of Ren YY who assisted in developing the algorithm and performing the data analysis, Tai CS who contributed towards study design, algorithm development and manuscript review, Syed Muhd Fahmy who assisted in retrieval and preparation of histological specimens, Diagnostics Development Hub of Exploit Technologies Pte. Ltd., Singapore for support in planning and executing the study and the physicians in the Department of Gastroenterology & Hepatology, Singapore General Hospital for contributing subjects for the study.

Author Contributions

Conceptualization: Pik Eu Chang, George Boon Bee Goh, Kiat Hon Lim, Chee Kiat Tan.

Data curation: Pik Eu Chang, George Boon Bee Goh, Wei Qiang Leow, Kiat Hon Lim, Chee Kiat Tan.

Formal analysis: Pik Eu Chang, George Boon Bee Goh, Wei Qiang Leow, Liang Shen, Kiat Hon Lim, Chee Kiat Tan.

Funding acquisition: Pik Eu Chang, Chee Kiat Tan.

Investigation: Pik Eu Chang, George Boon Bee Goh, Wei Qiang Leow, Chee Kiat Tan.

Methodology: Pik Eu Chang, George Boon Bee Goh, Wei Qiang Leow, Liang Shen, Kiat Hon Lim, Chee Kiat Tan.

Project administration: Pik Eu Chang, Kiat Hon Lim, Chee Kiat Tan.

Resources: Pik Eu Chang, Kiat Hon Lim, Chee Kiat Tan.

Supervision: Pik Eu Chang, Chee Kiat Tan.

Validation: Pik Eu Chang, Wei Qiang Leow, Liang Shen, Kiat Hon Lim.

Visualization: Pik Eu Chang.

Writing – original draft: Pik Eu Chang.

Writing – review & editing: Pik Eu Chang, George Boon Bee Goh, Wei Qiang Leow, Kiat Hon Lim, Chee Kiat Tan.

References

1. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. American Gastroenterological Association.; American Association for the Study of Liver Diseases.; American College of Gastroenterology. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology*. 2012 Jun; 142(7):1592–609. <https://doi.org/10.1053/j.gastro.2012.04.001> Epub 2012 May 15. PMID: 22656328
2. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol*. 2016 Jun; 64(6):1388–402. <https://doi.org/10.1016/j.jhep.2015.11.004> Epub 2016 Apr 7. PMID: 27062661
3. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 2011; 34:274–285. <https://doi.org/10.1111/j.1365-2036.2011.04724.x> PMID: 21623852
4. 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 Jun; 55(6):2005–23. <https://doi.org/10.1002/hep.25762> PMID: 22488764
5. Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology*. 2015 May; 61(5):1547–54. <https://doi.org/10.1002/hep.27368> PMID: 25125077.
6. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwithaya 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 Aug; 149(2):389–97.e10. <https://doi.org/10.1053/j.gastro.2015.04.043> PMID: 25935633.
7. Brunt EM. Nonalcoholic steatohepatitis. *Sem Liver Dis* 2004; 24(1):3–20
8. 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* (Baltimore, Md). 2011 Jun; 53(6):1874–82. <https://doi.org/10.1002/hep.24268> PMID: 21360720.

9. Castera L, Chan HL, Arrese M, Afdhal N, Bedossa P, Friedrich-Rust M, et al. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol*. 2015 Jul; 63(1):237–64. <https://doi.org/10.1016/j.jhep.2015.04.006> PMID: 25911335
10. Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003; 38:1449–1457. <https://doi.org/10.1016/j.jhep.2003.09.022> PMID: 14647056
11. Hui SL, Zhou XH. Evaluation of diagnostic tests without gold standards. *Statistical Methods in Medical Research*. 1998; 7:354–370. <https://doi.org/10.1177/096228029800700404> PMID: 9871952
12. 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(4):874–880. <https://doi.org/10.1002/hep.21346> PMID: 17006934
13. Jung ES, Lee K, Yu E, Kang YK, Cho MY, Kim JM, et al. Interobserver agreement on pathologic features of liver biopsy tissue in patients with nonalcoholic fatty liver disease. *Journal of pathology and translational medicine*. 2016; 50(3):190. <https://doi.org/10.4132/jptm.2016.03.01> PMID: 27086596
14. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for non-alcoholic fatty liver disease. *Hepatology* 2005; 41: 1313–1321 <https://doi.org/10.1002/hep.20701> PMID: 15915461
15. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999; 94: 2467–2474 <https://doi.org/10.1111/j.1572-0241.1999.01377.x> PMID: 10484010
16. Juluri R, Vuppalanchi R, Olson J, Ünalp A, Van Natta ML, Cummings OW, et al. Generalizability of the NASH CRN Histological Scoring System for Nonalcoholic Fatty Liver Disease. *J Clin Gastroenterol*. 2011 Jan; 45(1): 55–58
17. Friedman SL. Mechanisms of hepatic fibrogenesis. *Gastroenterology* 2008; 134: 1655–1669 <https://doi.org/10.1053/j.gastro.2008.03.003> PMID: 18471545
18. Bedossa P. Harmony in liver fibrosis. *J Hepatol* 2010, 52(3):313–314 <https://doi.org/10.1016/j.jhep.2009.11.020> PMID: 20133004
19. Sun W, Chang S, Tai DC, Tan N, Xiao G, Tang H, et al. Nonlinear optical microscopy: use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. *J Biomed Opt*. 2008; 13(6):064010. <https://doi.org/10.1117/1.3041159> PMID: 19123657
20. Tai DC, Tan N, Xu S, Kang CH, Chia SM, Cheng CL, et al. Fibro-C-Index: comprehensive, morphology-based quantification of liver fibrosis using second harmonic generation and two-photon microscopy. *J Biomed Opt*. 2009 Jul-Aug; 14(4):044013. <https://doi.org/10.1117/1.3183811> PMID: 19725725
21. Gailhouste L, Le Grand Y, Odin C, Guyader D, Turlin B, Ezan F, et al. Fibrillar collagen scoring by second harmonic microscopy: a new tool in the assessment of liver fibrosis. *J Hepatol*. 2010 Mar; 52(3):398–406. <https://doi.org/10.1016/j.jhep.2009.12.009> Epub 2010 Jan 17. PMID: 20149472
22. 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 Aug; 61(2):260–9. <https://doi.org/10.1016/j.jhep.2014.02.015> PMID: 24583249
23. Pirhonen J, Arola J, Sädevirta S, Luukkonen P, Karppinen SM, Pihlajaniemi T, et al. Continuous Grading of Early Fibrosis in NAFLD Using Label-Free Imaging: A Proof-of-Concept Study. *PLoS One*. 2016 Jan 25; 11(1):e0147804. <https://doi.org/10.1371/journal.pone.0147804> PMID: 26808140
24. Wang Y, Vincent R, Yang J, Asgharpour A, Liang X, Idowu MO, et al. Dual photon microscopy based quantitation of fibrosis-related parameters (q-FP) to model disease progression in steatohepatitis. *Hepatology*. 2017; 65(6): 1891–1903. <https://doi.org/10.1002/hep.29090> PMID: 28133774
25. Otsu N, A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems Man & Cybernetics*, 1979; 9(1):62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
26. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: Methodological review and comparison with logistic regression. *Annals of Behavioral Medicine*. 2003; 26(3):172–181. https://doi.org/10.1207/S15324796ABM2603_02 PMID: 14644693
27. Rückstieß T, Osendorfer C, van der Smagt P. Sequential Feature Selection for Classification. *AI 2011: Advances in Artificial Intelligence*, 2011; 7106:132–141. https://doi.org/10.1007/978-3-642-25832-9_14
28. Kearns K, Ron D, Algorithmic Stability and Sanity-Check Bounds for Leave-One-Out Cross-Validation. *Neural Computation*. 1999 Aug; 11(6): 1427–1453. <https://doi.org/10.1162/089976699300016304> PMID: 10423502