

SCIENTIFIC REPORTS



OPEN

High resolution Physio-chemical Tissue Analysis: Towards Non-invasive *In Vivo* Biopsy

Received: 20 January 2015
Accepted: 21 September 2015
Published: 04 February 2016

Guan Xu¹, Zhuo-xian Meng², Jian-die Lin², Cheri X. Deng³, Paul L. Carson¹, J. Brian Fowlkes¹, Chao Tao⁴, Xiaojun Liu⁴ & Xueding Wang^{1,3}

Conventional gold standard histopathologic diagnosis requires information of both high resolution structural and chemical changes in tissue. Providing optical information at ultrasonic resolution, photoacoustic (PA) technique could provide highly sensitive and highly accurate tissue characterization noninvasively in the authentic *in vivo* environment, offering a replacement for histopathology. A two-dimensional (2D) physio-chemical spectrogram (PCS) combining micrometer to centimeter morphology and chemical composition simultaneously can be generated for each biological sample with PA measurements at multiple optical wavelengths. This spectrogram presents a unique 2D “*physio-chemical signature*” for any specific type of tissue. Comprehensive analysis of PCS, termed PA physio-chemical analysis (PAPCA), can lead to very rich diagnostic information, including the contents of all relevant molecular and chemical components along with their corresponding histological microfeatures, comparable to those accessible by conventional histology. PAPCA could contribute to the diagnosis of many diseases involving diffusive patterns such as fatty liver.

The century old practice of staining biological tissue for examination under a microscope remains the standard assessment tool for pathologists looking for signs of diseases. Unfortunately, limited by the observation depth of optical microscopy, conventional pathological procedures rely on *ex vivo* tissue processed on slides, which require time-consuming preparation procedures, often accompanied by spatial distortion generated during sample fixation and staining. Innovative imaging technologies capable of characterizing the three-dimensional (3D) micron-scale morphology and functions of unfixed and unstained tissues *in vivo* may provide an alternative to the current histopathology and lead to revolutionary new tools for biology and medicine. Besides significant advantages such as less-invasiveness, lower cost and less time-lag to diagnosis, tissue characterization *in vivo* utilizing advanced imaging technologies could also improve sampling in the body, greatly benefiting diagnosis with higher sensitivity and specificity.

The accurate pathologic diagnosis of a disease is essentially achieved by evaluating both the microscopic morphological features, namely physical microstructure, and the chemical contents of tissue¹. Most currently available imaging technologies, however, are mono-physical and can only register either chemical components or physical microstructures in a biological tissue. Moreover, the findings from most of current imaging technologies are difficult to quantify and highly dependent on the systems and operators. Therefore, diagnostic imaging, in most cases, still cannot replace histology, nor provide similar information earlier. The noninvasive evaluation of the microstructures in deep tissue has been investigated extensively by using technologies such as quantitative-diffusion-tensor magnetic resonance

¹Department of Radiology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. ²Life Sciences Institute and Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. ³Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA. ⁴Key Laboratory of Modern Acoustics, Nanjing University, Nanjing, 210093, China. Correspondence and requests for materials should be addressed to X.W. (email: xdwang@umich.edu)

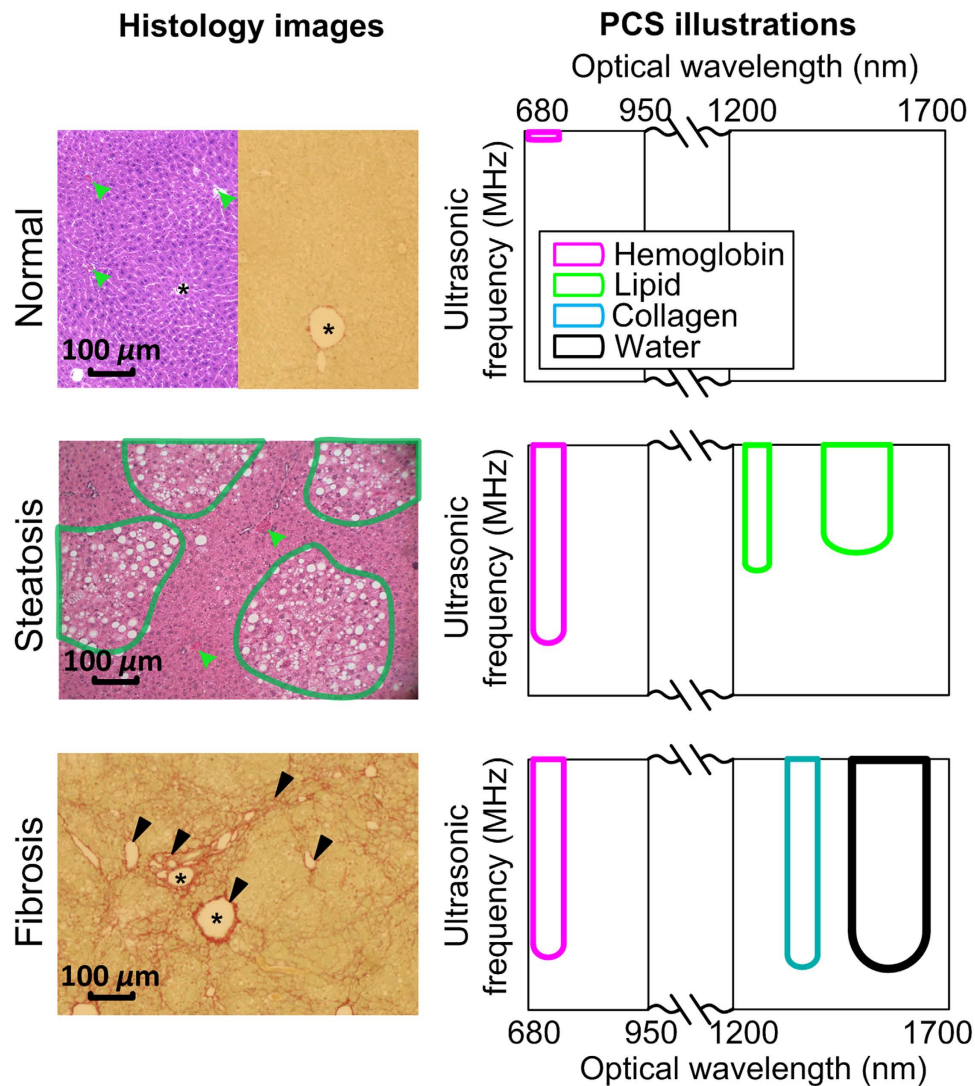


Figure 1. Illustration of the concept of PCS. Left column are histology photographs of normal, steatotic and fibrotic livers. Left half of the normal and steatosis histology photographs are H&E stained. The steatotic regions in the steatosis histology are outlined in green. Right half of the normal and fibrosis histology photographs are stained with Sirius Red. Red blood cells in the blood vessels and sinusoids are marked by green arrows. Fibrotic region is stained in red and marked by black arrows. In the histology photographs, “*” indicates the blood vessel or bile duct. As shown in the histology, abnormal liver conditions involve not only the changes in chemical components but also the changes in physical microstructures. Right column: the sketched PCS of the three liver conditions, including normal, steatosis and fibrosis, over the optical wavelengths covering strong absorption of hemoglobin, lipid, collagen and water. Each PCS shows distinctive signatures (i.e. combinations of the strips) which is correlated to the optically absorbing chemical components in the livers. Typically the heterogeneous distribution of a chemical component increases the extension of its corresponding fingerprint along the axis of US frequency. The concentration of a chemical component is correspondent to the magnitude of its PCS fingerprint, which will be demonstrated in Fig. 2.

imaging² and ultrasound (US) spectral analysis (USSA)³. Validated in a wide variety of organs, USSA has successfully correlated the dimensions and distributions of the US backscatters in biological tissues to the frequency-domain power distribution of the radio-frequency (RF) US signal^{4,5}. However, as an intrinsically mono-physics imaging modality, pulse-echo US imaging cannot interrogate the molecular components or chemical substances forming these backscattering microfeatures. Other modes of US offer further capabilities, but most offer much lower spatial resolution, as in shear wave elastography⁶. The specificity and sensitivity of the US imaging are thereby limited when identifying diseases with pathogenesis involving both microstructural and chemical changes. As a specific example, it is very difficult to differentiate liver fibrosis from steatosis based on US imaging. As both fat infiltration in hepatocytes [steatosis, as shown in the histology photograph in Fig. 1] and the collagen deposition in the

extracellular spaces [fibrosis, also shown in Fig. 1] exist in non-alcoholic fatty liver diseases (NAFLD), the US backscattering in liver may be caused by either fat or collagen clusters⁷.

Conventional spectroscopic optical imaging (SOI) technologies provide approaches for evaluating the chemical components in biological tissues by measuring their intrinsic optical absorption spectra^{8,9}. By recognizing the unique optical absorption spectrum of each chemical material, SOI over a broad spectrum could achieve quantitative chemical imaging with satisfactory sensitivity. However, as a result of the overwhelming light scattering in biological tissues, conventional SOI has only limited imaging depth. Moreover, the limited spatial resolution of SOI makes it challenging to recover histological microfeatures in subsurface tissues.

Optically (photo-) induced ultrasonic (-acoustic) imaging, namely photoacoustic (PA) imaging (PAI), is an emerging nonionizing and nonradiative imaging modality combining the advantages of both optical and US imaging^{10–12}. In PAI, the tissue sample is illuminated by a pulsed laser. When absorbed by biological tissue, the optical energy will convert to heat and cause thermoexpansion in the tissue. The thermoexpansion leads to the vibrations of the sample which generate propagating waves at a broad-band of frequencies, namely PA waves. The PA waves within the ultrasonic range can be captured by US transducers for later image reconstruction. PAI thereby enables the imaging of the optical properties in deep biological tissue at ultrasonic resolution.

Encoded in the high frequency PA signal is the spatial distribution of optical absorption properties in biological tissue. PAI thereby allows for the extension and combination of optical absorption spectroscopy and microscopic tissue architecture characterization in a single, eventually portable diagnostic system. Based on the PA measurements of each tissue specimen, a two-dimensional (2D) physio-chemical spectrogram (PCS) which integrates the power spectra of RF PA signals along the full optical spectrum can be formulated. One axis of the PCS in optical wavelengths represents the optical absorption spectrum of the tissue; the other in ultrasonic frequencies represents the spatial frequencies of the statistically repeating and optically absorbing structures in the tissue. The unique PCS of a tissue sample thereby encodes rich diagnostic information comparable to that in pathology, including the chemical components within the tissue along with their corresponding physical microstructures. Comprehensive and objective assessment of tissues, as well as diagnosis and treatment monitoring of a variety of diseases could be achieved by the analysis of PCS, namely PA physio-chemical analysis (PAPCA), without invasive and complicated biopsy procedures.

In this study, we have examined the feasibility of PAPCA in tissue characterization and its potential to realize the desired non-invasive *in vivo* biopsy through the studies on established mouse models of liver diseases. *Ex vivo* and *in situ* experiments were conducted on mice with extreme liver steatosis and fibrosis. In addition, PCS of mouse livers with progressive NAFLD, including the conditions of steatosis and fibrosis, were characterized by multivariate analysis methods to examine the feasibility of PAPCA in diagnosis and characterizing the severity of NAFLD.

Results

The laboratory animal protocol for this work was approved by the University Committee on Use and Care of Animals of the University of Michigan. All the animal experiments in this study were carried out in accordance with the approved guidelines.

Concept of PCS. Figure 1 schematically demonstrated the concept of the PCS. Similar to SOI, the optical absorption spectrum, or the “color”, of each liver specimen representing the composition of various chemical components can be observed by the PA signal amplitudes at different optical wavelengths. By putting together the power spectra of PA signals acquired at all the optical wavelengths over the entire optical-spectrum-of-interest, a PCS of a target tissue is formulated. The microstructures in the liver tissues correlate to the ultrasonic frequency domain power distribution of the PA signals, i.e., the heterogeneous distribution of the chemical components produces high ultrasonic frequency components in the PA signals. With more high ultrasonic frequency components, the power spectra appear more extended along the ultrasonic frequency axis, as shown by the vertical stripes in the 2D PCS in Fig. 1, right column. The stripe features, namely PCS fingerprints, along the ultrasonic frequency are expected at the optical absorption peaks of the major chemical components of the tissue sample. Due to the differences in chemical components and microstructures, each liver condition should possess a unique 2D PCS signature as a combination of the PCS fingerprints.

Study on extreme models of liver steatosis and fibrosis. *PCS of normal, extreme steatotic and extreme fibrotic livers.* Figure 2 shows the typical PCS of three different liver conditions, including normal control, extreme steatosis and extreme fibrosis, obtained through the *ex vivo* measurements on mouse livers. At each optical wavelength in the PCS in Fig. 2, by summing up the spectral amplitudes in normal scale from 0.1 MHz to 8 MHz [where the signal-to-noise ratio (SNR) is approximately 1], the optical absorption spectrum of each tissue sample can be obtained, as shown row 4 in Fig. 2. By looking at the unique optical spectrum of each absorbing material in row 5 in Fig. 2, one can understand the chemical contents of the three types of liver samples as combinations of hemoglobin, lipid, collagen and water.

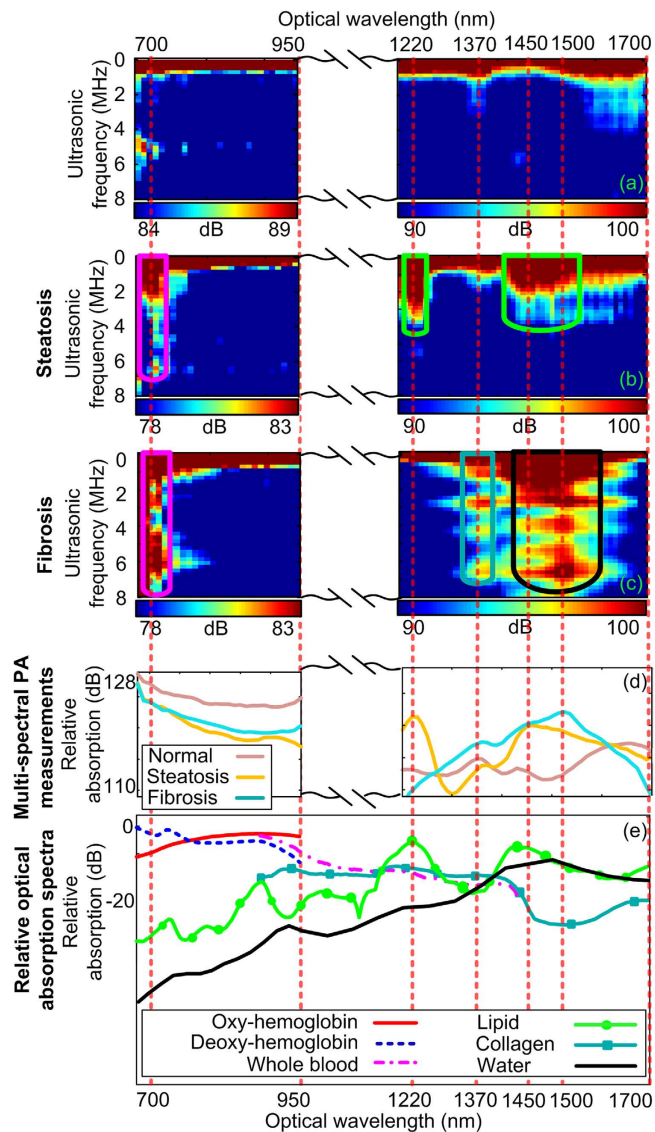


Figure 2. Example PCS spectrograms of normal, extreme steatosis, extreme fibrotic livers. The fingerprints of hemoglobin, lipid, collagen and water are marked by the magenta, green, blue and black contours, respectively. The dynamic range of the normal liver PCS in row 1 is maintained identical to those of the other two PCS and is adjusted to show the quick decrease of spectral magnitude along the ultrasonic frequency axis at 700 nm, i.e. less extended hemoglobin fingerprint. The multi-spectral PA measurements of normal, extreme steatosis, and extreme fibrotic livers are shown in row 4. The curves in row 4 are produced by summing the pixels in the PCS maps in normal scale along each column. The relative optical absorption spectra of the major optically absorbing chemical components in liver tissue are shown in row 5^{29–32}. The red dashed lines indicate the matching between the absorption peaks of the chemical components and the fingerprints in the PCS.

In comparison with the normal control, the distinctive PCS of the steatotic liver demonstrates: 1) a stronger lipid fingerprint in 1200–1240 nm due to the increased lipid content in the steatotic liver (i.e. chemical change), and 2) an extended lipid fingerprint toward higher ultrasonic frequency which is a result of scattered and heterogeneous deposition of the lipid in the steatotic liver (i.e. physical micro-structure change). Both findings are confirmed by the histology images in Fig. 1. Besides the changes in lipid fingerprint, the PCS of the steatotic liver also shows differences at the fingerprint associated with hemoglobin at 700 nm. Lower overall spectral magnitude (Fig. 2, row 4) and fingerprints extended into the high ultrasonic frequency (Fig. 2, row 2) are found in the PCS of the steatotic liver sample compared to that of the normal liver sample in Fig. 2. This is because the sinusoids distorted by lipid accumulation become insufficient conduits of blood, leading to the reduction of hemoglobin per microscopic field within the affected volumes in liver^{13,14}. The distribution of such affected area within the liver volume

formulated the heterogeneous distribution of hemoglobin in steatotic livers. Similar histological changes have been reported in fibrotic livers¹⁵.

By comparing the PCS of fibrotic liver with that of the normal control, the differences at the collagen fingerprint in 1350–1390 nm can be observed, including 1) the higher magnitude indicating the higher collagen content, and 2) further extension of the fingerprint suggesting the enhanced heterogeneity in collagen distribution in the fibrotic liver. Both of the observations are resulted from the deposition of collagen in the extracellular spaces in liver as confirmed by the histology in Fig. 1. Similar to the situation in the steatotic liver, the fibrotic liver also shows lower magnitude and further extension at the hemoglobin fingerprint around 700 nm (Fig. 2), which is due to the decrease of hemoglobin content and the heterogeneous hemoglobin distribution caused by collagen deposition.

The PCS of the steatotic and the fibrotic livers also show broad fingerprints, within the optical spectral range of 1400–1600 nm. A close observation reveals the different mechanisms causing these changes compared to the normal liver. In the PCS of the steatotic liver, the broad fingerprint peaks at 1450 nm which, according to the optical absorption spectra in row 5 in Fig. 2, is another absorption peak of lipid. Compared to the total spectral magnitude of the other two liver conditions, the fibrotic liver has larger spectral magnitude at 1370 nm and 1400–1500 nm in Fig. 2 row 4. The strong optical absorption at 1370 nm correlates to the strong optical absorption of collagen, as shown in the optical absorption spectra of the chemical components in Fig. 2 row 5. Since the fibrotic liver sample does not contain lipid deposition, the strong optical absorption within 1400–1500 nm of fibrotic is associated with the increase of water content, which possesses strong absorption in the range of 1400–1600 nm, as shown in Fig. 2 row 5. However, due to the partial overlapping of the optical absorption spectra of lipid and water, the spectral information in 1400–1600 nm will not be considered for liver tissue characterization in this study.

Non-invasive in situ imaging of mouse livers. The differences in PCS among different liver conditions raise the possibility that the lipid and the collagen fingerprints in the PCS could be the deterministic features to characterize extreme steatotic and fibrotic livers. The identification of these two liver conditions was attempted by *in situ* and non-invasive imaging experiments in mice. Since blood content is not of major concern in this experiment and the use of live animal involves strong motion artifacts, the mice were sacrificed right before the imaging experiment. The setup and procedures are described in the Methods section. PA signals were acquired at 1220 and 1370 nm for quantifying the lipid and collagen contents in comparison to the normal controls, respectively. The power spectra of the RF PA signals, i.e. the vertical lines in PCS at 1220 and 1370 nm, were quantified by a previously developed method, PASA¹⁶. The A-lines of the beamformed PA signals were segmented to formulate localized PASA parameter images.

A commercially available linear array (L7-4, Philips Healthcare, Andover, MA) was used to cover the ultrasonic frequency range below 10 MHz, which is sufficient to distinguish the liver area from the background tissues and the hundred-micron level microfeatures in liver steatosis and fibrosis shown in histology (Fig. 1). Figure 3(a) shows the result from an obese mouse with severe liver steatosis compared to that from a normal control. In both US and PA intensity images, no obvious difference between the two liver conditions can be recognized. Comparing to the normal liver tissue, the steatotic liver shows prominently different PASA parameters, as demonstrated by each of the pseudo-color 2D PASA parameter images (intercept, midband-fit, and slope, respectively.) superimposed on the PA intensity images. The increase of lipid content in steatotic liver tissue leads to the increase of its spectral amplitude and, consequently, the PASA parameters including intercept and midband-fit; while the heterogeneous distribution of the lipid infiltrated hepatocytes increased the extension of the lipid fingerprint which is quantified as larger slope values. The contrasts in the spectral magnitudes (especially the intercept) between the normal and steatotic livers are not as prominent as those in the spectral slope. This is due to the fact that the varied thickness of the subcutaneous fat in the obese and normal mice brings in uncertainty to the attenuation of the light energy. Compared to intercept and midband-fit, the slope is less susceptible to the light fluence and, hence, is more reliable for quantitative imaging and objective tissue characterization.

The US gray-scale, PA intensity and PASA parameter images of a normal and a fibrotic mouse liver are compared in Fig. 3(b). Due to the deposition of collagen fibers and, consequently, the increase of tissue heterogeneity, the fibrotic liver shows prominent contrast in PASA parameters over the normal one at 1370 nm, including higher intercept, higher midband-fit, and higher slope.

The finding from this imaging experiment suggests that, in comparison to conventional PA intensity imaging, PASA at the optical absorption peaks of relevant chemical materials, by quantitatively and objectively assessing tissue histological microstructures, may render higher contrast when presenting different liver conditions. The extreme steatosis model, as demonstrated in this study, only involves the change of lipid content in liver; whereas the extreme fibrosis model only involves the change of collagen content. In this case, these two liver conditions can be identified by PASA of a single vertical line in PCS corresponding to the strong optical absorption wavelength of lipid or collagen, respectively.

Study on a progressive NAFLD model. *PCS of progressive NAFLD liver at different stages.* Besides the extreme cases, NAFLD in realistic scenarios involves advanced stages where steatosis and fibrosis co-exist in the liver. The non-invasive, accurate and objective grading instead of simply identifying the

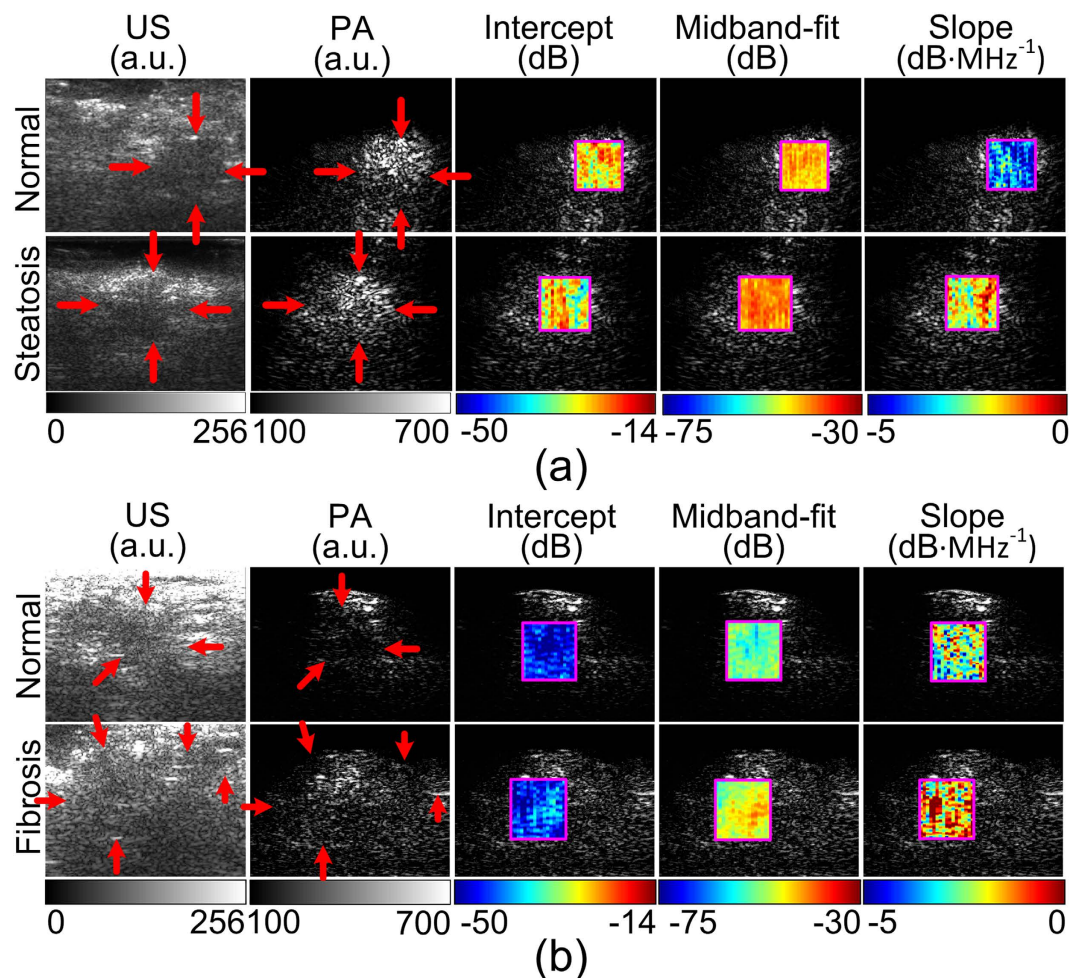


Figure 3. Non-invasive US and PA imaging and PASA of normal, extreme steatosis, and extreme fibrotic mouse livers *in situ*. The red arrows indicate the boundaries of the livers. The unit of the image dimensions is mm. (a) Normal liver vs steatotic liver at 1220 nm optical wavelength. In comparison with the normal control, the steatotic liver shows higher intercept, higher midband-fit and higher slope values. (b) Normal liver vs fibrotic liver at 1370 nm optical wavelength. In comparison with the normal control, the fibrotic liver shows higher intercept, higher midband fit, and higher slope values.

liver conditions has thereby greater clinical value. To further examine the sensitivity of our method to the pathological changes in liver, a mouse model (STAM, Stelic Institute and Co., Tokyo, Japan) with quick yet full spectrum of NAFLD progression was employed. The PCS of a total of 36 mouse liver specimens were acquired, including 12 normal, 12 steatotic and 12 fibrotic liver specimens. As shown by Fig. 4, the representative PCS of the three liver conditions in this progressive NAFLD mouse model demonstrate findings similar to those from the extreme steatotic and fibrotic mouse livers, including:

- (1) Both PCS of the fibrotic and steatotic livers show extended hemoglobin fingerprints at 700 nm;
- (2) Steatotic livers demonstrate enhanced and extended lipid fingerprints at 1220 nm and 1450 nm. The fibrotic livers also show discernible lipid fingerprints, corresponding to the remaining lipid contents shown in the histology images.
- (3) Fibrotic livers demonstrate enhanced and extended collagen fingerprints at 1370 nm. The steatotic livers also show noticeable yet less distinctive collagen fingerprints, generated by the small amount of collagen content shown in the histology photograph in Fig. 4.

Categorization of the liver conditions by Support Vector Machine (SVM). Unlike the situation in the extreme steatotic and fibrotic liver models where the histological changes involve either lipid or collagen, as shown in Fig. 1, the concurrence of steatosis and fibrosis in the different stages of the progressive NAFLD model makes these stages difficult to be distinguished. Moreover, the animal-to-animal difference further increases the challenge to establish unanimous criteria for staging NAFLD by PASA parameter at a single wavelength. Figure 5 shows the statistics of the experiment data from the progressive

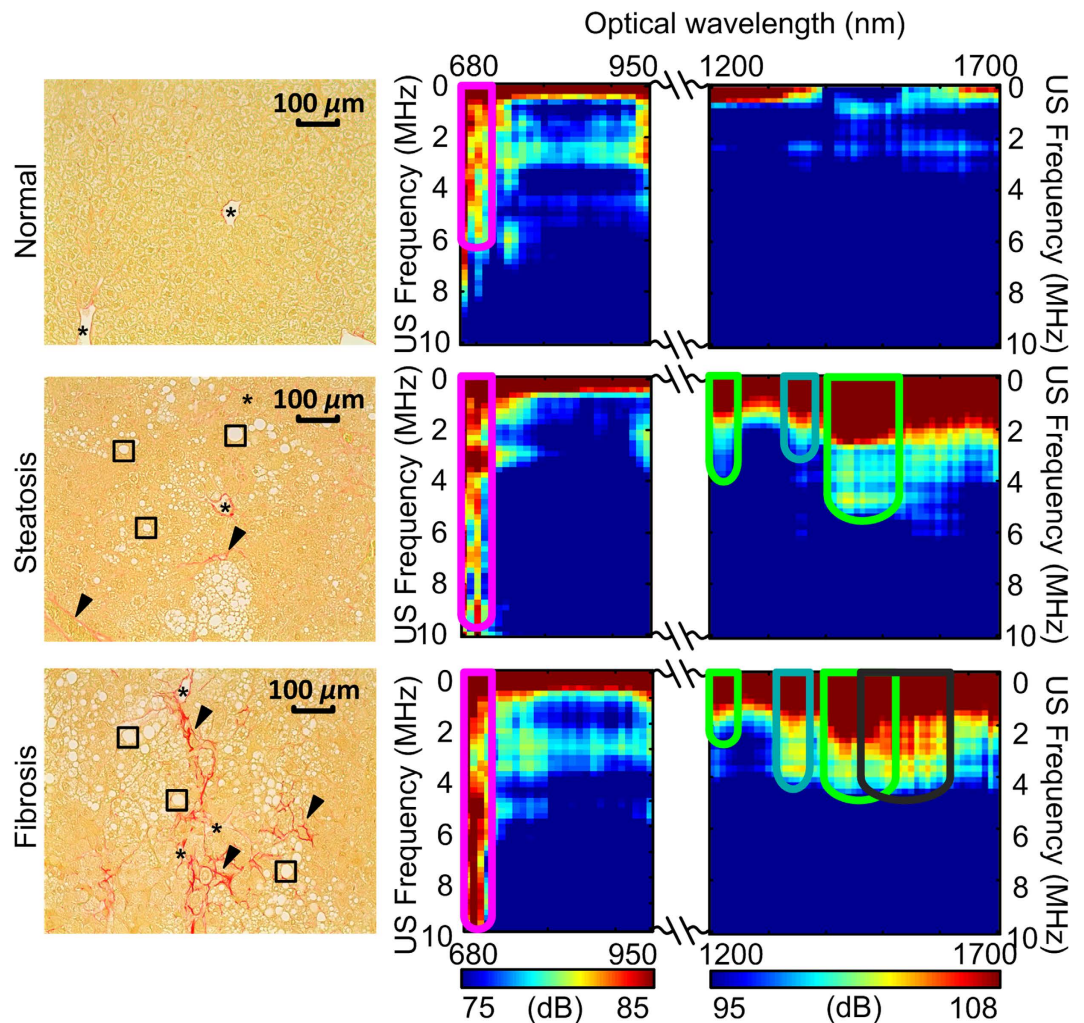


Figure 4. The histology photographs and PCS of the normal, steatosis and fibrosis stages in a progressive NAFLD model. In the histology photographs, “*” indicate the blood vessel or bile ducts; “□” indicate lipid infiltrated hepatocytes; and black arrows indicate collagen content. Following Fig. 2, the fingerprints of hemoglobin, lipid, collagen and water are marked by the magenta, green, blue and black contours, respectively.

NAFLD mice, where the slope and the midband-fit at 700 nm, 1220 nm and 1370 nm are used independently to differentiate the three liver conditions. The average values of the PASA parameters agree well with the expected trends associated with hemoglobin, lipid and collagen changes. For example, livers at steatotic stage have high lipid content and strong heterogeneity in lipid distribution; while livers at fibrotic stage have high collagen content and strong heterogeneity in collagen distribution. Although the three groups were distinguishable by 25 to 75 percentile of the data distribution, the large deviations of and the overlaps between the groups make it difficult to accurately categorize the three groups using single PASA parameters. As examples, neither the slope nor the midband-fit values at 1220 nm can differentiate the liver steatosis and fibrosis; and the midband-fit values at 1370 nm cannot separate steatosis and fibrosis samples.

PAPCA possesses the advantage of simultaneously quantifying the absolute and relative changes of a variety of chemical components in their spatial heterogeneity and concentrations. This study characterizes the physio-chemical information by implementing a multi-variant analysis tool, SVM¹⁷ to categorizing the PASA parameters at multiple optical wavelengths-of-interest, including 700 nm, 1220 nm and 1370 nm. A 3-fold cross-validation approach¹⁸ was used in this study to avoid overfitting the SVM model and to fully understand the performance of categorizing the liver conditions by SVM. The 12 sets of data acquired from each liver condition were randomly divided into 3 groups with 4 data sets. 2 of the 3 groups for each liver conditions were used in turns for training the SVM and the rest will be used for testing, that is, a total of 24 data sets (2 groups × 4 data sets × 3 conditions) for training and 12 data sets (1 group × 4 data sets × 3 conditions) for testing. As shown in Table 1, such arrangement led to three training/testing cycles for categorization with either of slope and midband-fit or both. The

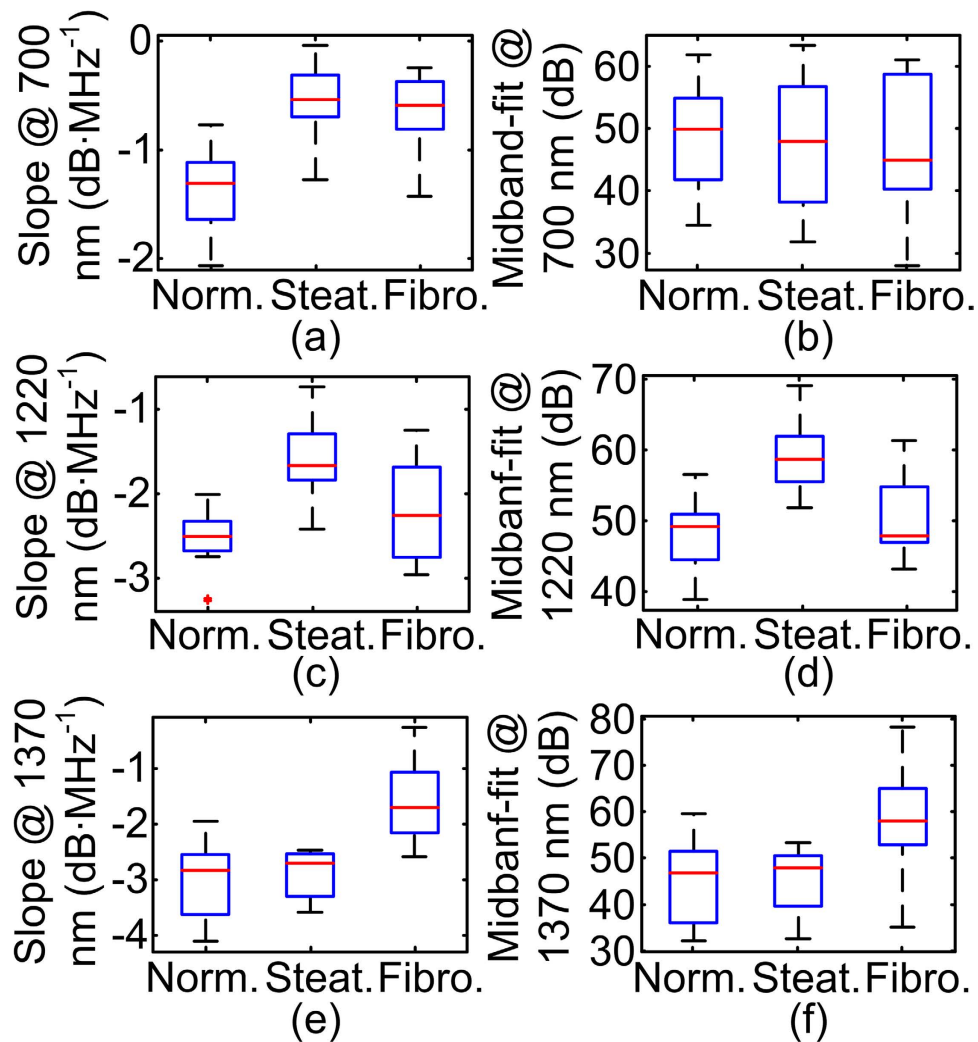


Figure 5. Statistics of the experiment data from the progressive NAFLD mice. Each data group consists of 12 data points. The red lines represent the averages of the data points. The upper and lower edges of the boxes are the 25th and 75th percentiles, respectively. The dashed lines extend to the most extreme data points and do not consider outliers. The outliers are plotted as “+”.

training and testing accuracies were averaged to evaluate the SVM categorization performance, as shown in Table 2. Figure 6(a,b) show the scattered plot of the data points and the SVM models determined in the first training cycle with slopes and midbandfits, respectively. The formulation of the SVM models is further described in the method section. As shown in Table 2, The SVM trained using only the slopes and only the midband-fits can achieve averaged accuracy of 83.3% and 50%, respectively in categorizing the test data sets. Nonetheless, when both spectral parameters are used, the accuracy can be improved to 88.9%. Such results substantiate our hypothesis that the spectral magnitude can be less reliable than the slope values due to the unpredictable nature of light energy distribution in optically-scattering biological tissues.

Discussion

The study on the mouse models of NAFLD has demonstrated that the PCS from the PA measurement of a biological tissue contains very rich diagnostic information. Taking advantage of the endogenous contrasts of the chemical components in liver, PCS has presented not only the contents of a variety of chemical components but also the histological microstructures associated with these chemical components. Each type and condition of tissues has unique “signature” in the 2D PCS which also shows satisfying repeatability. For example, in Figs 2 and 4, steatotic or fibrotic livers induced by different pathogenesis show similar PCS. The non-invasive *in situ* experiments in Fig. 3 have also lead to the same conclusion as those in our previous study¹⁹. In future clinical setting, systematic analysis of the PCS, i.e. PAPCA, holds the potential for noninvasively detection and objective grading of progressive NAFLD by comprehensively evaluating both physical and chemical changes in liver tissues *in vivo*. Besides liver steatosis

Test sample condition	Cate. #	Test samp.#	Cycle 1			Cycle 2			Cycle 3		
			Slope c=3 $\gamma=1/6$	Mid c=10 $\gamma=1/9$	Both c=3 $\gamma=1/2$	Slope c=5 $\gamma=1/8$	Mid c=10 $\gamma=1/4$	Both c=60 $\gamma=1/2$	Slope c=10 $\gamma=1/6$	Mid c=10 $\gamma=1/6$	Both c=100 $\gamma=1/5$
Normal	1	1	1	1	1	1	1	1	1	3	1
	1	2	1	2	1	2	1	1	1	2	1
	1	3	1	1	1	2	1	1	1	1	1
	1	4	1	1	1	1	1	1	1	1	1
Steatosis	2	5	2	2	2	2	1	1	2	2	2
	2	6	2	2	2	2	3	2	2	1	2
	2	7	2	2	2	2	1	2	2	1	2
	2	8	2	3	2	2	3	2	1	1	1
Fibrosis	3	9	3	3	3	3	3	3	3	3	3
	3	10	1	3	3	3	1	3	3	1	3
	3	11	3	1	3	3	1	3	1	1	1
	3	12	3	2	3	3	3	3	2	1	1
Training accuracy			95.8% (23/24)	100% (22/24)	100% (24/24)	95.8% (23/24)	100% (24/24)	95.8% (23/24)	100% (24/24)	100% (24/24)	100% (24/24)
Categorization accuracy			91.7% (11/12)	66.7% (8/12)	100% (12/12)	83.3% (10/12)	50% (6/12)	91.7% (11/12)	75% (9/12)	33.3% (4/12)	75% (9/12)

Table 1. SVM categorization. A 3-fold cross-validation has been performed by using all the data sets in turns for training and testing the SVM model. The data sets are randomly divided into 3 groups. Each cycle uses 2 groups for training and 1 group for testing c is capacity constant and γ is the parameter in Eq. (1) for the polynomial kernel. Cate. = category. samp. = sample.

and fibrosis, the same method could also be extended to the diagnosis of other NAFLD related conditions such as steatohepatitis and cirrhosis. The characteristic change in steatohepatitis is the formation

	Slope	Mid	all
Mean Training accuracy	97.2%	97.2%	98.6%
Mean Training accuracy	83.3%	50%	88.9%

Table 2. SVM categorization accuracy. The numbers are derived by averaging the training and testing accuracies with the same data combination for all 3 cycles in Table 1.

of macrovesicular and inflammatory foci, composed predominantly of inflammation cells (including lymphocytes)²⁰. Such focal structure could be captured in PCS by implementing fluorescence agents targeted at the lymphocytes²¹. The cirrhosis induces the further loss of hemoglobin²², the escape of the lipid droplets²³ and the further deposition of the collagen in liver¹⁵, all of which can be reflected by the change of the corresponding fingerprints in PCS. Although PAPCA was first validated through the study on liver conditions, the same technology could be adapted to the diagnosis of many other diseases including, but not limited to, kidney diseases, cardiovascular diseases or cancer where diffusive patterns of the chemical components can be found and are crucial for evaluating the severity of the diseases⁵. The diagnostic value of PAPCA could be further augmented by introducing optical contrast enhancing agents.

The liver tissues were scanned *ex vivo* and *in situ* right after the animals are sacrificed. Such approach minimizes the motion artifacts in the measurements and data averaging for optimal SNR. However, the disadvantage of using sacrificed animal is the inability to observe oxy- and deoxygenated hemoglobin individually. 700 nm illumination is thereby selected for measuring total hemoglobin optical absorption with sufficient penetration depth. In the *ex vivo* experiments, the optical and acoustic attenuation caused by the acoustic coupling materials was maintained the same for each scan by cutting the sample into same thickness. The attenuations were afterwards calibrated before the formation of the PCS following the methods described in the method section. However, for the *in situ* case, the variation of the skin and adipose thickness could bring in uncertain optical as well as ultrasonic attenuations to the measurements. Strong optical and ultrasonic attenuation could lead to low SNR, which leads to less prominent contrast in midband-fit values in Fig. 5. On the other hand, as shown in the experiment results, the slope is less affected as it describes the relative descending trend of the spectral magnitudes along the frequencies. Even though the part of the PCS could be overwhelmed by noise level, the descending trend

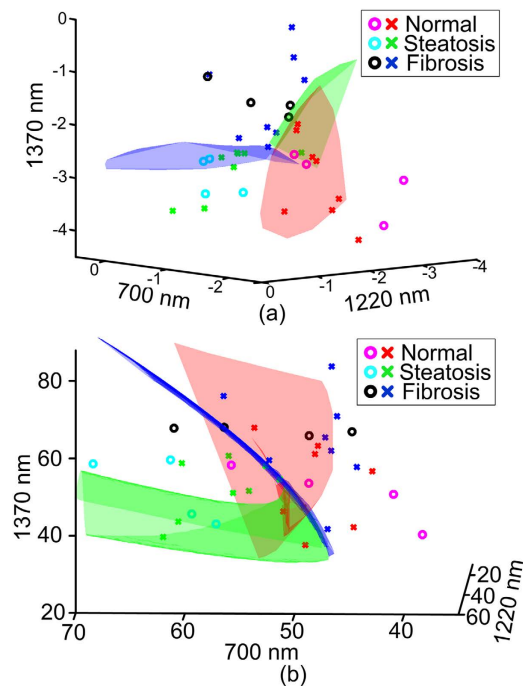


Figure 6. SVM categorization of the data in Fig. 5. (a) SVM categorization based on slopes at 700, 1220 and 1370 nm. (b) SVM categorization based on midband-fits at 700, 1220 and 1370 nm. “o” are training data and “x” are testing data.

of the spectra above the noise level could still carry microstructural information in the livers. Since the mouse skin is relatively thin, sufficient SNR has been achieved in this study. The stiffness of the livers, which could be higher in the fibrotic livers compared to the other two conditions²⁴, could also bring in sample-wise variations of ultrasonic frequency dependent attenuation to the PA signals. However, since the sampling area is relatively small and the ultrasonic frequency range observed is relatively low, the stiffness of the samples was not considered and will be investigated in future studies. The optical and ultrasonic attenuation in the PA measurements will be significant in human subjects. Future studies will also include compensating these attenuations using the thickness of skin and adipose tissue in US images.

The PCS from liver tissues are characterized by performing PASA at three specific optical wavelengths corresponding to one dominant optical absorption chemical components in the liver. The PASA method facilitates quick and quantitative evaluation of both amplitudes and extensions of the PCS fingerprints, which represent the concentrations and the histological microfeatures of the relevant chemical components in the tissue samples. The PASA is performed on the linear fitting of the averaged power spectra of the RF PA signals, which provides a cogent means of addressing the stochastic nature of tissue microstructure and leads to measurements that are quantitative and repeatable. However, the linear approximation approach in the PASA does not consider the fluctuations in the signal power spectrum. Similar to the findings in USSA, the periodically fluctuating patterns in the power spectra could be another representation of the cluster sizes²⁵ or repetitive distributions of the optical absorbers. The loss of varied amount of blood during the sample harvesting leads to the large uncertainty in the optical fluence within the samples and subsequently PA spectral magnitudes around 700 nm. Large variation of midband-fits in Fig. 5(b) is thereby observed. However, slopes describe the relative descending trend of PA spectra and are less affected by the blood content variation as long as sufficient SNR is achieved. The PCS fingerprints in the wavelength range of 1400–1500 nm were not included as these fingerprints contain the content and microscopic structural information of both the lipid and water. Solving for the chemical components involves decoupling both optical and ultrasonic information in the 2D PCS, which will be investigated in future studies.

The SVM used in this study is fundamentally a binary classifier, which could involve ambiguous decision²⁶ when extended to multi-classification tasks. 3-fold cross-validation which uses all the data sets for training and testing in turns is used to avoid SVM overfitting and biased results. The parameter selection for optimal SVM categorization was achieved by grid search described in the user guide to the SVM package²⁷. Other techniques such as feature selection may be investigated in future study for more effective SVM model optimization. SVM categorization based on neither slope nor midband-fit can achieve the accuracy when both parameters are used. This is because slope and midband-fit values are independent and provide orthogonal diagnostic information. As shown in the training/testing cycle 1 in Table 1, using only slopes, the SVM cannot correctly categorize sample 10. On the other hand, SVM

trained by midband-fits successfully identified the conditions of sample 10. The slopes and midband-fits thereby compensated each other in SVM categorization using both parameters and achieved 100% accuracy. Similar observation can be found for sample 2 and 3 in cycle 2 in Table 1. The testing accuracies in most cases are lower than the training accuracies. Therefore, although 100% accuracy in cycle 1 in Table 1 has been achieved using both slopes and midband-fits, as the sample size increases, the categorization accuracy could decrease due to the individual differences and measurement noises.

For the progressive NAFLD liver model, the observation at the limited number of isolated optical wavelengths is sufficient for identifying the three stages. The relative changes among specific chemical components were not considered. Compared to conventional optical or ultrasound technologies, PAPCA possesses the unique advantage of being able to simultaneously quantifying the absolute and relative changes of a variety of chemical components in their spatial heterogeneity and concentrations which may lead to more accurate evaluation of the disease conditions. For example, the ratios between the extensions (i.e. slopes) of different fingerprints associated with various relevant chemical substances may also contain diagnostic information. Further efforts will thereby be dedicated to extracting the comprehensive diagnostic information included in the PCS with more advanced analysis tools by recognizing the entire 2D “signature” in the PCS and evaluating the contribution of each parameter to the diagnosis.

Methods

Mouse models. The extreme steatotic liver used for acquiring the PCS in Fig. 2 was generated in a C57BL/6J wild type mouse from Jackson laboratory. The obese mouse was fed with chow diet for the first 8 weeks, followed by 60% high-fat diet (Research diet, D12492) for the 12 weeks thereafter. The control mouse was fed with chow diet for 20 weeks. Extreme steatosis in the experiment group with no fibrosis was confirmed by histology in Fig. 1.

The extreme fibrosis model used for the PCS spectrum in Fig. 2 was FIP200 liver knock out (LKO) mouse, which spontaneously develops fibrosis without steatosis²⁸. The mouse was fed with chow diet for 15 weeks. As is shown in the histology images in Fig. 1, extreme fibrosis yet no steatosis has occurred.

Both extreme models were designed to demonstrate the most distinctive features of the one liver condition in PCS without the influence of the other. For both conditions, C57BL/6J mice were used as control group.

For the categorization study, we used a mouse model which quickly develops the progressive NAFLD in liver (STAM, Stelic Institute & Co, Tokyo, Japan). As confirmed by the histological images in Fig. 4, the mouse livers formulate substantially observable steatosis at week 6 and fibrosis at week 12, although the steatosis and fibrosis level were not as high as those in Fig. 1. A total of 36 mouse liver specimens, 12 from each category of normal, steatotic and fibrotic livers, were examined in the *ex vivo* experiments.

PCS scanning system. The data acquisition systems in this study are shown in Fig. 7. The optical illumination in the system was generated by a tunable Optical Parametric Oscillator (OPO) laser (Vibrant B, Oportek Inc, Carlsbad, CA, USA) pumped by the second harmonic output of an Nd:YAG pulsed laser (Brilliant B, Quantel, Bozeman, MT, USA) with spectral linewidth of 15–20 cm⁻¹. The tuning ranges of the laser are 680–950 nm and 1200–1700 nm, covering the strong absorption range of hemoglobin and the vibrational absorption peaks of lipid, collagen and water in near-infrared (NIR) spectrum^{29–32}. A constant tuning step size of 10 nm was set during the PCS acquisition. The energy output over the entire tuning range was controlled to 15–20 mJ per pulse, with a pulse duration of 8 ns. The laser beam was collimated to 12.5 mm in diameter, leading to an incident light fluence of 11–15 mJ/cm² which is below the American National Standards Institute (ANSI) safety level. 10% of the laser energy was projected to an optical power meter by a beam splitter for energy monitoring and later signal magnitude calibration.

For *ex vivo* studies in Fig. 7(a), the received PA signals were collected by a needle hydrophone (HNC-1500, ONDA Co., Sunnyvale, CA, USA) with a receiving bandwidth of 0–10 MHz and sensitivity of −242 dB re 1V/μPa. The location of the sample holder as well as the hydrophone, and the receiving direction of the hydrophone are all fixed to assure minimized measurement variability. The signals received by the hydrophone was amplified and then digitized by a digital oscilloscope (TDS 540, Tektronix, Inc., Beaverton, OR, USA) before being collected by a PC. In the *in situ* experiments in Fig. 7(b), the data acquisition was conducted with our PA-US dual-modality real-time imaging system¹⁰. PA signals were acquired using a commercial US transducer array (L7-4, Philips, Andover, MA) covering an ultrasonic frequency band below 10 MHz. For all data acquisition, the PA signals were averaged 30 times to improve SNR.

PCS formulation and PASA. The PA signals from the tissue generated at each wavelength, after being transformed into ultrasonic frequency domain, were calibrated by the laser energy, the frequency dependent attenuation of the acoustic signals and the system frequency-dependent response. In the *ex vivo* study, the acoustic attenuation was estimated by considering the acoustic path as the distance between the hydrophone surface and the center of the sample¹⁶. The quantitative attenuation was calculated using the data in a previous study³³. The frequency response of the hydrophone was provided by the vendor. In the *in situ* study, since the US probe was attached close to the liver tissue, the acoustic attenuation cause by the skin and adipose tissue was less significant and thereby not considered. The frequency response of the US probe system was quantified by measuring the PA signals from a hair

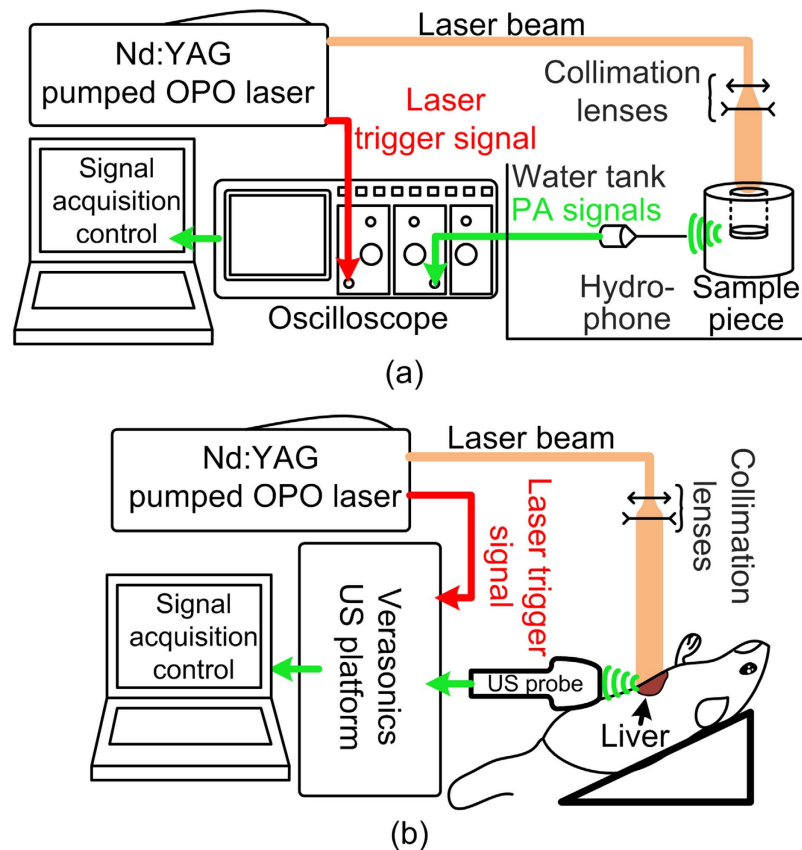


Figure 7. Experiment setups. (a) Setup for imaging of *ex vivo* liver specimens. Both the hydrophone and the sample piece were submerged in water for acoustic coupling. The laser beam was expanded by the collimation lenses to 12.5 mm in diameter and projected to the sample. The sample was held by a tube made of porcine gel (Sigma-Aldrich, St. Louis, MO). Such design minimizes the attenuation of the optical illumination by coupling water. (b) Setup for noninvasive imaging of mouse liver *in situ* in a model cadaver. The expanded laser beam was projected onto the mouse liver through the abdomen wall. US gel was used for acoustic coupling.

fiber³⁴, which has the dimension ($30\ \mu\text{m}$) beyond the spatial resolution of the US probe and can be regarded as an impulse to the data acquisition system. The complete set of power spectra from all the optical wavelengths formed a PCS, as shown in Figs 2 and 4, with the optical wavelengths shown along the horizontal axis, the US frequencies shown along the vertical axis, and the intensities of the power spectra shown in pseudo-color.

For quantitative and objective analysis of PCS to achieve tissue characterization, the PA signal power spectra at the strong optical absorbing wavelengths of relevant chemical substances (e.g. hemoglobin, lipid, and collagen) were analyzed by the PASA methods as described in our previous works¹⁶. To be brief, the power spectrum of the PA signal from the target tissue was first fit to a linear model. Then, three spectral parameters, including intercept, midband-fit and slope, of the linear model were quantified. Among the three, only two are independent. Since both intercept and midband-fit represent the spectral magnitude, and midband-fit appears more robust to the uncertainties in measurements in Fig. 3, intercept is not included in the SVM categorization in Table 1.

Statistical study and categorization with SVM. As mentioned in 4.1, each of the 36 mouse liver specimens was scanned, and a total of 36 PCS were generated. With each PCS, the slopes and midband-fit values at the three optical wavelengths of 700 nm, 1220 nm and 1370 nm were quantified. The mean and the standard deviation of the slopes and midband-fits within each liver condition group (i.e. normal, steatosis, and fibrosis) were calculated by the built-in functions in MATLAB (2011b, Mathworks, Boston, MA) and shown in Fig. 5. A multi-variant analysis MATLAB tool, LIBSVM³⁵ was used to examine the feasibility of characterizing the liver conditions using the PASA parameters (i.e. the slope and the midband-fit at the three optical wavelengths including 700 nm, 1220 nm and 1370 nm). The PASA parameters from 8 liver specimens in each category (normal, steatosis, and fibrosis) were used to train the SVM. The PASA parameters from the rest 4 specimens in each category were used for examining

the diagnosis accuracy. We used a type 1 (namely C-type in LIBSVM) SVM with 3rd order homogeneous polynomial kernels^{17,35}. The polynomial kernel is defined as:

$$K(x_i, x_j) = (\gamma x_i^T x_j)^3 \quad (1)$$

where x is the variable matrix, i and j are the indices of the matrix. γ is an adjustable parameter specific for each application. The categorization was first based on either slope or midband-fit. Afterwards, both parameters were used for improved categorization accuracy. When both parameters were used, the slope and midband-fit values were normalized within [0, 1] respectively. The capacity constant and the γ in the polynomial kernel were selected for each categorization case by attempting with exponentially increasing values following the user guide to LIBSVM³⁵ and shown in Table 1. As previously mentioned, 3-fold cross-validation approach¹⁸ was used in this study to avoid overfitting the SVM model and to fully understand the performance of categorizing the liver conditions by SVM. The training and testing accuracies were averaged to demonstrate the SVM categorization accuracy, as shown in Table 2.

References

- Kumar, V., Abbas, A. K., Fausto, N. & Aster, J. C. *Robbins & Cotran pathologic basis of disease*. (Elsevier Health Sciences, 2009).
- Basser, P. J. & Pierpaoli, C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson* **213**(2), 560–570 (2011).
- Lizzi, F. L. *et al.* Theoretical framework for spectrum analysis in ultrasonic tissue characterization. *J Acoust Soc Am* **73** (4), 1366–1373 (1983).
- Hall, T. J., Insana, M. F., Harrison, L. A. & Cox, G. G. Ultrasonic measurement of glomerular diameters in normal adult humans. *Ultrasound Med Biol* **22** (8), 987–997 (1996).
- Lizzi, F. L., Feleppa, E. J., Kaisar Alam, S. & Deng, C. X. Ultrasonic spectrum analysis for tissue evaluation. *Pattern Recogn Lett* **24** (4–5), 637–658 (2003).
- Sebag, F. *et al.* Shear wave elastography: a new ultrasound imaging mode for the differential diagnosis of benign and malignant thyroid nodules. *J Clin Endocr Metab* **95** (12), 5281–5288 (2010).
- Saverymattu, S. H., Joseph, A. E. & Maxwell, J. D. Ultrasound scanning in the detection of hepatic fibrosis and steatosis. *Brit Med J* **292** (6512), 13–15 (1986).
- Xu, G., Piao, D. & Dehghani, H. The Utility of Direct-Current as Compared to Frequency Domain Measurements in Spectrally-Constrained Diffuse Optical Tomography Toward Cancer Imaging. *Tech cancer Res Treat* **10** (5), 403–416 (2011).
- Pogue, B. W. *et al.* Quantitative Hemoglobin Tomography with Diffuse Near-Infrared Spectroscopy: Pilot Results in the Breast. *Radiology* **218** (1), 261–266 (2001).
- Yuan, J. *et al.* Real-time photoacoustic and ultrasound dual-modality imaging system facilitated with graphics processing unit and code parallel optimization. *J Biomed Opt* **86001**, 1 (2013).
- Wang, X. *et al.* Noninvasive laser-induced photoacoustic tomography for structural and functional *in vivo* imaging of the brain. *Nat Biotech* **21** (7), 803–806 (2003).
- Wang, L. V. Multiscale photoacoustic microscopy and computed tomography. *Nat Photonics* **3** (9), 503–509 (2009).
- Ijaz, S., Yang, W., Winslet, M. C. & Seifalian, A. M. Impairment of Hepatic Microcirculation in Fatty Liver. *Microcirculation* **10** (6), 447–456 (2003).
- Farrell, G. C., Teoh, N. C. & McCuskey, R. S. Hepatic Microcirculation in Fatty Liver Disease. *The Anat Rec* **291** (6), 684–692 (2008).
- Bataller, R. & Brenner, D. A. Liver fibrosis. *J Clin Invest* **115** (2), 209–218 (2005).
- Xu, G. *et al.* Photoacoustic spectrum analysis for microstructure characterization in biological tissue: A feasibility study. *App Phys Lett* **101** (22), 221102–221105 (2012).
- Cortes, C. & Vapnik, V. Support vector machine. *Mach learn* **20** (3), 273–297 (1995).
- Varma, S. & Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* **7** (1), 91 (2006).
- Xu, G. *et al.* The Functional Pitch of an Organ: Quantification of Tissue Texture with Photoacoustic Spectrum Analysis. *Radiology*, **271** (1), 248–254 (2014).
- Brunt, E. M. *et al.* Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroen* **94**, 2467–2474 (1999).
- Parish, C. R. Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunol Cell Biol* **77** (6), 499–508 (1999).
- Bradley, S. E., Ingelfinger, F. & Bradley, G. Hepatic cirrhosis in cirrhosis of the liver. *Circulation* **5** (3), 419–429 (1952).
- Hartroft, W. S. & Ridout, J. H. Pathogenesis of the Cirrhosis Produced by Choline Deficiency: Escape of Lipid from Fatty Hepatic Cysts into the Biliary and Vascular Systems. *Am J Pathol* **27**, 951 (1951).
- Ziol, M. *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. *Hepatology* **41**, 48–54 (2005).
- Lizzi, F. L. *et al.* Relationship of Ultrasonic Spectral Parameters to Features of Tissue Microstructure. *IEEE T Ultrason Ferr* **34**, 319–329 (1987).
- Chih-Wei, H. & Chih-Jen, L. A comparison of methods for multiclass support vector machines. *IEEE T Neural Networ* **13** (2), 415–425 (2002).
- Lin, C. J., Hsu, C.-W. & Chang, C.-C. A practical guide to support vector classification. *National Taiwan U.* www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf (2003) Date of access 05/01/2015.
- Ma, D. *et al.* Autophagy Deficiency by Hepatic FIP200 Deletion Uncouples Steatosis From Liver Injury in NAFLD. *Mol Endocrin* **27** (10), 1643–1654 (2013).
- Cheong, W. F., Prahl, S. A. & Welch, A. J. A review of the optical properties of biological tissues. *IEEE J Quantum Elect* **26** (12), 2166–2185 (1990).
- Jacques, S. L. Optical properties of biological tissues: a review. *Phys Med Biol* **58** (11), R37 (2013).
- Querry, M. R., Cary, P. G. & Waring, R. C. Split-pulse laser method for measuring attenuation coefficients of transparent liquids: application to deionized filtered water in the visible region. *App Opt* **17** (22), 3587–3592 (1978).
- Wang, H.-W. *et al.* Label-Free Bond-Selective Imaging by Listening to Vibrationally Excited Molecules. *Phys Rev Lett* **106** (23), 238106 (2011).
- Pinkerton, J. M. M. The Absorption of Ultrasonic Waves in Liquids and its Relation to Molecular Constitution. *Proc Phys Soc B* **62** (2), 129 (1949).

34. Kumon, R. E., Deng, C. X. & Wang, X. Frequency-Domain Analysis of Photoacoustic Imaging Data From Prostate Adenocarcinoma Tumors in a Murine Model. *Ultrasound Med Biol* **37** (5), 834–839 (2011).
35. Chang, C.-C. & Lin, C.-J. LIBSVM: a library for support vector machines. *ACM Trans Intel Sys Tech* **2** (3), 27 (2011).

Acknowledgements

We thank Dr. Bin Nan at the School of Public Health and Dr. Jieping Ye at Department of Computational Medicine and Bioinformatics in University of Michigan for their advice on the SVM approaches. This work was supported by the American Heart Association grant 14POST17840001, Michigan Institute for Clinical & Health Research UL1TR000433 and National Institute of Health under grant numbers R01AR060350 and R01CA186769.

Author Contributions

G.X. wrote the main manuscript and prepared the figures. G.X., C.X.D., P.L.C., J.B.F., C.T., X.L. and X.W. contributed to the development of the PA and US imaging technologies in this study. Z.X.M. and J.D.L. contributed to the animal model preparation and histology of the liver tissues. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xu, G. *et al.* High resolution Physio-chemical Tissue Analysis: Towards Non-invasive *In Vivo* Biopsy. *Sci. Rep.* **6**, 16937; doi: 10.1038/srep16937 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>