

# H-scan trajectories indicate the progression of specific diseases

Jihye Baek | Kevin J. Parker

Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York, USA

## Correspondence

Kevin J. Parker, University of Rochester, Rochester, NY, USA.  
Email: kevin.parker@rochester.edu

## Funding information

National Institutes of Health, Grant/Award Number: R21EB025290

## Abstract

**Purpose:** The ability of ultrasound to assess pathology is increasing with the development of quantitative parameters. Among these are a set of parameters derived from the recent H-scan analysis of subresolvable scattering. The emergence of these quantitative measures of tissue/ultrasound interactions now enables a study of the unique trajectories of multiparametric features in multidimensional space, representing the progression of specific diseases over time. We develop the mathematical and visual tools that are effective for classifying, quantifying, and visualizing the steady progression of several diseases from independent studies, all within a uniform framework.

**Methods:** After applying the H-scan analysis of ultrasound echoes, we trained a support vector machine (SVM) to classify the unique trajectories of progressive liver disease from fibrosis, steatosis, and pancreatic ductal adenocarcinoma (PDAC) metastasis. Our approaches include the development of trajectory maps and disease-specific color imaging stains.

**Results:** The multidimensional SVM image classification reached 100% accuracy across the three different studies.

**Conclusion:** H-scan trajectories can be useful to track the progression of multiple classes of diseases, improving diagnosis, staging, and assessing the response to therapy.

## KEYWORDS

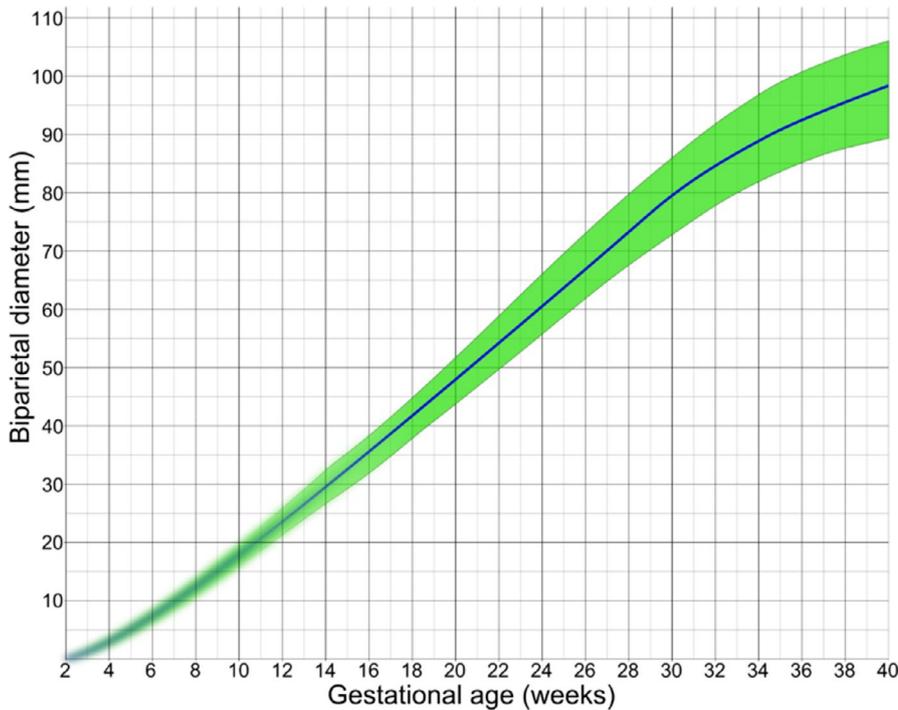
fibrosis, metastases, multiparametric ultrasound, steatosis, support vector machine

## 1 | INTRODUCTION

The value of tracking ultrasound anatomical measurements over time has been enshrined for decades in the chart of biparietal diameter vs. gestational age, a widely used reference for monitoring the health and growth of the fetus.<sup>1-3</sup> An example is shown in Figure 1.

Over the years, the list of useful anatomical measurements has grown. These are structures that are large enough to be resolved by medical ultrasound scanners and identified over time to assess changes.

Examples include the cross sectional area of nerves,<sup>4</sup> cardiovascular and arterial measures,<sup>5,6</sup> muscle disease,<sup>7</sup> kidney cysts,<sup>8</sup> and cancer volumes.<sup>9</sup> However, the reproducibility of anatomical measures of 3D anatomy with ultrasound has been limited by user factors and the predominance of 2D imaging systems.<sup>8,9</sup> Separately, Doppler ultrasound measurements over time have seen useful applications in a range of conditions, including rheumatoid arthritis, fetal development,<sup>10</sup> and cardiovascular diseases.<sup>11,12</sup> The majority and the simplest of these measurements involve a



**FIGURE 1** A plot of fetal development represented by ultrasound measurement of cranial diameter as a function of time during pregnancy. The blue line represents the mean and the green area represents the 90% prediction interval. By Mikael Häggström, used with permission<sup>3</sup> [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

single parameter measured at two time points, creating a delta or change (difference) measure. At a higher level of complexity lies the ability to estimate multiple parameters over time.<sup>13-16</sup> When extended to subresolvable scattering properties, this enables an assessment of the measured parameters that provide the highest discrimination between groups: for example responders and non-responders to chemotherapy<sup>17</sup>; normal and abnormal liver conditions.<sup>18-20</sup>

We propose that the diagnostic value of multiparametric ultrasound measures can be advanced by combining the following concepts:

1. by applying the H-scan analysis, which is linked to the scattering and wave propagation physics within tissues;
2. by creating trajectories of H-scan (and other) measures in multidimensional spaces, representing the progression of disease or response to treatments; and
3. by deriving color-coded images based on proximity to clusters in multidimensional parameter spaces, representing distinct pathological conditions.

In effect, the original concept of the biparietal diameter curve is extended to a higher dimensional space, but with important transformations. First, the simple anatomical measure of size is replaced by advanced metrics related to the biophysics of scattering and wave propagation, representing interactions that occur on a subresolvable scale within conventional imaging systems. Second, the progression of disease or pathology must be mapped out in the multidimensional parameter space and segmented into unique pathways or clusters.

Finally, the tools for visualizing, simplifying, and segmenting multiparameter spaces are needed to reduce the dimensionality and make practical visualizations of the results. Among these tools, we find that principal component analysis (PCA) and the support vector machine (SVM) are robust and practicable means to achieve a clinically useful trajectory that can be visualized and used as a diagnostic guide. The importance of the final representation of the measured data cannot be understated; as Tufte says "...certain methods for displaying and analyzing data are more likely to produce truthful, credible, and precise findings. The difference between an excellent analysis and a faulty one can sometimes have momentous consequences."<sup>21</sup>

## 2 | MATERIALS AND METHODS

In order to classify, quantify, and visualize the steady progression of several diseases from independent studies within a uniform framework, the following experiments were conducted and analyzed.

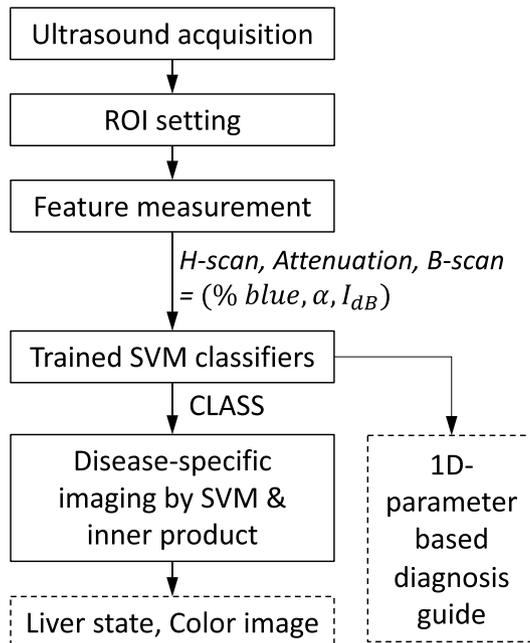
### 2.1 | Animal models

We investigated three unique diseases of pancreatic ductal adenocarcinoma (PDAC) liver metastasis, fibrosis, and steatosis from separate *in vivo* studies. These animal studies are summarized in Figure 2(a). The PDAC liver metastasis<sup>22,23</sup> was developed in 9 mice (C57BL/6J) that were purchased from Jackson Laboratory (Bar Harbor, ME, USA). To grow the tumor in livers,  $4 \times 10^5$  luciferase expressing murine pancreatic

(a)

Disease	Enrolled animals	Ultrasound scanner	Transmission	$f_c$ (Transducer)	Acquired Frames
Steatosis (MCD)	12	Visualsonics, vevo 3100	Focused	15 MHz (MX 201)	395
Fibrosis (CCl <sub>4</sub> )	11	Visualsonics, vevo 2100	Focused	21 MHz (MS 250)	1062
Tumor (PDAC)	9	Verasonics, Vantage 256	Plane wave	10 MHz (L11-5v)	1337

(b)



### c. Disease-specific imaging

For each pixel,

$$v = (H, A, B) \\ = (\% \text{ blue}, \alpha, I_{AB}) \\ - (50 \%, 0.5 \text{ dB/MHz/cm}, -15 \text{ dB})$$

**switch** CLASS

**case** Normal; Color = gray scale B-scan

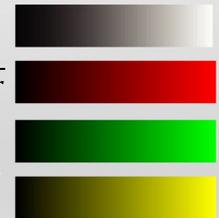
**case** Tumor; Color =  $[1 \ 0 \ 0] \times \frac{\langle v, e_1 \rangle}{\max \text{Tumor}}$

**case** Fibrosis; Color =  $[0 \ 1 \ 0] \times \frac{\langle v, e_2 \rangle}{\max \text{Fib}}$

**case** Steatosis; Color =  $[1 \ 1 \ 0] \times \frac{\langle v, e_3 \rangle}{\max \text{Fat}}$

**end**

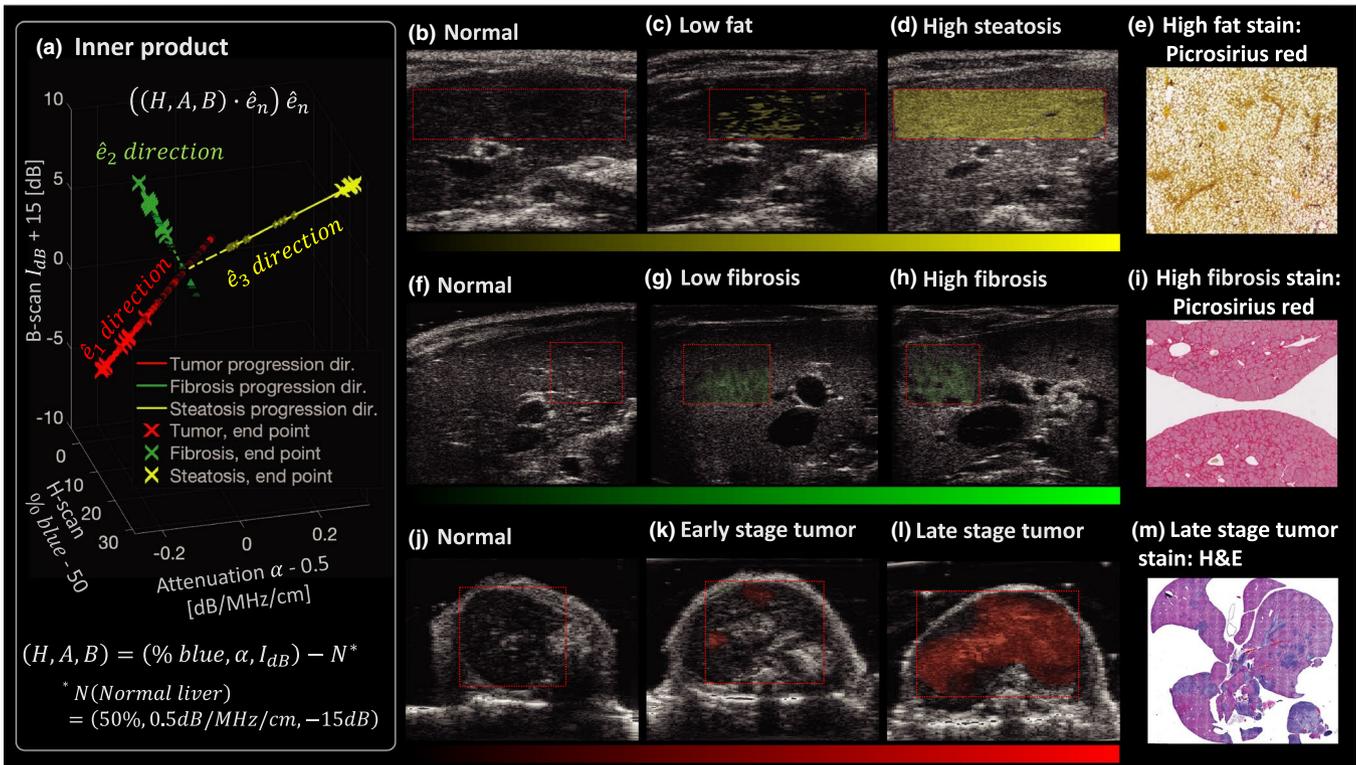
Display = B-scan  $\times$  0.5 + Color  $\times$  0.5



**FIGURE 2** (a) Study design and ultrasound acquisition are summarized. The three disease models were investigated and the number of enrolled animals for each model were provided. Three ultrasound scanners scanned the animals using the transmission beam types with their center frequencies ( $f_c$ ). Each animal was scanned repeatedly over time to monitor disease progression, and for each scan, several frames were acquired, the total number of acquired frames are provided. (b) Flow chart suggesting liver diagnosis guide, including *disease-specific imaging* and 1D parameter method. First, clinicians would scan a liver and set ROI box. Then features are extracted using H-scan, attenuation estimation, and B-scan procedures. Using the measured features, trained SVM identify liver conditions for all pixels within the ROI, and disease-specific imaging displays color images. Furthermore, using the features, 1D parameter analysis by trained SVM is also performed to assist diagnosis based on H-scan trajectories in Figure 5. (c) Detailed color processing for *disease-specific imaging*. Input for the processing is features and CLASS from the SVM, which are obtained for all pixels within a ROI. A color for each pixel is assigned according to the determined CLASS, and then inner product between feature vector  $v$  and unit vector  $e_n$  for disease progression direction of the CLASS is performed to calculate color intensity. The unit vectors were also calculated when training SVM using all acquired frames in (a). The colors for PDAC tumor, fibrosis, and steatosis were set as red, green, and yellow, respectively. In this way, the color image is made and transparently overlaid on the B-scan, whereas the pixels identified as normal do not have colors, showing only B-scan [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

tumor cells (KCKO-luc) were injected into the spleen, and then the tumor cells spread through the hepatic veins, causing metastasis in the liver. The mice were divided into two groups of control ( $n = 4$ ) and treated ( $n = 5$ ). The treated group underwent chemotherapy with Gemcitabine (50 mg/kg dosing, twice per week). The tumor growth was confirmed by pathology and bioluminescence imaging (BLI). Hematoxylin and eosin (H&E) stained liver sections, provided different colors for normal and tumor cells as shown in Figure 3(m). As evidenced by BLI, all 9 enrolled mice showed an increase in radiance from D-Luciferin as the tumor grows. The PDAC study was approved by the University of Rochester Committee on Animal Resources.

The fibrosis was induced in response to carbon tetrachloride (CCl<sub>4</sub>) exposure (1mL/kg, three times per week) under approval of the Pfizer Animal Care and Use Committee. Eleven rats were enrolled: 7 Sprague-Dawley (SD, Charles River Laboratories, Wilmington, MA, USA) and 4 TAC NIH-RNU (nude, Taconic Biosciences, Inc., Rensselaer, NY, USA). After the CCl<sub>4</sub> dosing, an expert pathologist assessed fibrosis response using trichrome stain and Picrosirius red, and all the rats were graded as stage 4 fibrosis. However, since CCl<sub>4</sub> can also cause fat accumulation, we measured the fat area percentage by using Oil Red O stain, resulting in  $2.9 \pm 1.9\%$ . To verify the enrolled rats have fibrosis, but minimal fat accumulation, fat area for the



**FIGURE 3** Disease-specific imaging. (a) 3 features of liver tissues were extracted by H-scan, attenuation estimation, and B-scan, generating the components of  $(\% \text{ blue}, \alpha, I_{dB})$ , respectively. The coordinate represents  $(H, A, B) = (\% \text{ blue}, \alpha, I_{dB}) - (50, 0.5, -15)$  for convenience, meaning changes from normal liver (50%, 0.5 dB / MHz / cm, -15 dB). Linear lines show disease progression pathways from normal to end time point causing death. All measured features of this study were projected onto their diseases' pathway, showing scatter plots. The end points causing death are denoted using the symbol "×". Then the distance between the origin and the measurement can be used for color intensity, meaning the greater values represent more severe disease. SVM classifies liver states, and assigns colors of yellow, green, and red for steatosis, fibrosis, and tumor, respectively. Then, the colors are overlaid on B-scan (b-d, f-h, j-l). Normal in (b, f, j) only shows B-scan without colors. Early stages in (c, g, k) have some color area with low intensity. Late stages in (d, h, l) fill in with saturated colors, representing extensive disease. Pathology slides (e, i, m) show different cellular patterns for the three diseases, supporting the different colors for the three diseases [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

other 9 normal SD rats was also measured, resulting in  $1.2 \pm 1.5\%$ :  $P$ -value comparing the normal and fibrosis group is 0.05.

Regarding the steatosis, a methionine and choline deficient (MCD) diet was employed to activate a non-alcoholic fatty liver disease (NAFLD) model. Twelve SD rats (Charles River Laboratories, Wilmington, MA, USA) were fed the MCD diet (MP Biomedicals, Solon, OH) for 6 weeks. The fat accumulation was assessed by a pathologist using PicroSirius red and H&E. The steatosis study was approved by the Institutional Animal Care and Use Committee at the University of Texas at Dallas.

## 2.2 | Ultrasound data acquisition and experimental set up

Ultrasound scans were performed independently for the three disease models by using three different ultrasound systems; consequently, a total of 2778 ultrasound images were acquired as described in Figure 2(a).

First, a Verasonics ultrasound scanner (Vantage 256, VeraSonics, Inc., Kirkland, WA, USA) equipped with an L11-5v probe at a 10 MHz center frequency imaged the livers with progression of the tumor. Plane wave transmission (25 angles,  $-6^\circ$  to  $6^\circ$ ) generated IQ data in the Verasonics system. For all of the 9 enrolled mice for the PDAC model, the data acquisition was performed twice a week from day 11 after the tumor cell injection until the death of each mouse; the last scan days range from day 28 to day 64. At each time point for the same mouse, 3–5 frames were obtained, resulting in 395 radiofrequency (RF) datapoints. The fibrosis study used a Vevo 2100 (FUJIFILM VisualSonics, Toronto, ON, Canada) with a 21 MHz center frequency linear array transducer (MS 250). Focused beam transmission was used to acquire RF data once a week from baseline before dosing; the rats survived until 6 or 8 weeks. At each time point for the same rat, approximately 30 frames were obtained, and therefore 1062 scans were acquired. The ultrasound scanner for the steatosis model is Vevo 3100 (FUJIFILM VisualSonics, Toronto, ON, Canada) utilizing a 15 MHz center frequency linear probe (MX 201). RF

data acquisition with focused beam transmission was performed at baseline (before starting the MCD diet), week 2, and week 6. For each scan, approximately 30 frames were examined, resulting in 1337 scans.

Using the acquired ultrasound signals, we obtained B-mode images, and then set a ROI for each animal, including a consistent liver area over time for the same animal; example ROIs are provided in Figure 3(b–d, f–h, j–l) as the red boxes. The data within each ROI were assigned as inputs for H-scan, attenuation estimation, B-scan, and disease-specific imaging.

## 2.3 | Feature measurement

For each liver, the H-scan *%blue*, attenuation coefficient  $\alpha$  (dB/MHz/cm), and B-scan intensity  $I_{dB}$  (dB) were measured using the H-scan matched filter analysis. In general terms, these parameters pertain to the frequency-dependent transfer function of the tissue scatterers, the frequency-dependent attenuation coefficient of the tissue, and the relative scattered intensity, respectively. Specifically, the H-scan *%blue* reveals differences in transfer functions of scatterers and is implemented using convolution between ultrasound echo  $r(t)$  and a matched filter set, including  $n$  Gaussian filters  $\left\{ G_k \mid \exp\left(-\frac{f-GC_k}{2\sigma_G^2}\right) \text{ for } k = 1, \dots, n \right\}$  with the Gaussian

center frequencies  $GC_k$ , but the same band width  $\sigma_G$ . For each time point  $t$ ,  $r(t) * \mathcal{F}^{-1}(G_k)$  yields  $n$  convolution values, that are likely to have a single maximum, indicating any shift in spectrum of the echo signal  $r(t)$ . Consequently, H-scan analysis can find the spectral shift at each sampling point of RF data by finding the maximum across the convolution results. However, ultrasound propagation causes frequency- and depth-dependent attenuation, resulting in the frequency downshift along with depth. The frequency shift was used to estimate  $\alpha$ ,

$$\alpha(x) = -\frac{f_p(x) - f_c}{x\sigma^2}, \quad (1)$$

where  $f_p(x)$  is measured peak frequency (MHz) at depth  $x$  (cm).  $f_c$  and  $\sigma$  are the center frequency and bandwidth of received echo, respectively, but before attenuation effects; these can be considered as initial parameters. With the estimated  $\alpha$ , attenuation correction was performed, producing an attenuation corrected signal  $r'(t)$ . Next, the convolution  $r'(t) * \mathcal{F}^{-1}(G_k)$  achieves matched filter outputs and corresponding peak frequencies for the points  $\in (\text{Sample}, \text{Scanline})$  again, and pseudocolor<sup>24</sup> maps the frequency to H-scan color. Since each Gaussian filter has corresponding colors, there are  $n$  color levels ranging from red to blue. This study used  $n = 256$ , indicating there are 256 color levels; the lower 128 levels from 1 to

128 in sequence represent red colors from vivid to darker, respectively, whereas the higher 128 levels from 129 to 256 represent from dark blue to bright blue. Within the ROI of RF data, *%blue* is defined by

$$\%blue = \frac{(\text{number of blue pixels})}{(\text{number of total pixels})} \times 100\%. \quad (2)$$

More details for H-scan analysis including attenuation correction and estimation are found in Parker and Baek,<sup>25</sup> Baek et al.,<sup>16</sup> respectively. The *%blue* varies according to the Gaussian filter parameters of  $GC_k$  and  $\sigma_G$ , and therefore we needed to set the parameters with a consistent standard for the three different ultrasound scanners. For the consistent setting, a bandwidth ( $BW$ ) is defined at 0.3 of normalized frequency spectrum  $S(f)$ . Each scanner's  $\{GC_k \mid k = 1 \dots 256\}$  of Gaussian filters was set to have *%blue* = 50% for normal livers, and  $GC_{k=256} - GC_{k=1} = 0.7 \times BW$ .

The three ultrasound scanners have different procedures to obtain log-compressed B-mode data, and the B-mode intensity was generally optimized for each scanner. Therefore, for use of  $I_{dB}$  as a feature in addition to *%blue* and  $\alpha$ , the intensity of normal livers was set to  $-15$  dB.

After measuring the features as described above using the 2778 ultrasound scans, the features at the same time point for the same animal were averaged, and therefore we finally obtained 115, 47, and 36 data set for PDAC tumor, fibrosis, and steatosis, respectively, meaning that a total of the 198 averaged data (*%blue*,  $\alpha$ ,  $I_{dB}$ ) were used for all feature analyses.

## 2.4 | Non-linear parameter transformation suggesting a combined parameter

We proposed a non-linear transformation to produce a 1D parameter by combining two properties of PCA and the features obtained by H-scan and attenuation estimation. We kept the form of transformation  $Z = XW$  from PCA where  $X$  and  $Z$  are the input and output of PCA, respectively, and  $W$  is a weight matrix, but we determined the  $W$  based on the H-scan and attenuation results from livers. Normal livers are reported to have approximately 0.5 dB/MHz/cm attenuation and have 50% blue for H-scan measurement. When considering  $(\%blue, \alpha) \in \mathbb{R}^2$ , measurements from normal livers form clusters around (50, 0.5), and progression of any diseases makes the features  $(\%blue, \alpha)$  change from the normal values, following a pathway that is unique to the three studied diseases. The parameter changes of  $\Delta(\%blue)$  and  $\Delta\alpha$  indicate the progressions; these two parameters are independently changed according to disease types.

Prior to parameter transformation, including PCA, data normalization is commonly performed such as Z-score or min-max normalization.<sup>26</sup> To set zero-mean and unit-standard deviation, the normalized parameters  $x_H$  and  $x_A$  for %blue and  $\alpha$ , respectively, are obtained by

$$x_H = \frac{(\%blue) - 50}{10}; x_A = (\alpha - 0.5) \times 10, \quad (3)$$

where 50 and 0.5 are the averages of % blue and  $\alpha$ , respectively, setting the zero-mean. The scaling factor of 10 serves to set approximately unit-standard deviation and also comparable scales for the different measurements  $x_H$  and  $x_A$ . Considering PCA with normalized parameters of  $x_H$  and  $x_A$ , the first principal component  $PC_1$  is

$$PC_1 = w_1 \cdot x_H + w_2 \cdot x_A, \quad (4)$$

where  $w_1$  and  $w_2$  can be positive or negative and are optimized by PCA to maximize the variation of  $PC_1$ . However, after examining the way that PCA derives  $PC_1$  in Equation (4) and the particular direction of changes with disease progression, we propose a *Combined Parameter C*:

$$C = w_H \cdot |x_H| + w_A \cdot |x_A|, \quad (5)$$

where  $w_A = 0.6 \cdot \text{sign}(x_A)$  and  $w_H = 0.4 \cdot \text{sign}(x_H)$ . Note that both weights rely on the sign of attenuation, making this a nonlinear combination. The weights, 0.4 and 0.6, are optimized contributions from H-scan and attenuation, respectively, to maximize classification accuracy instead of the variance. Note that  $w_H$  and  $w_A$  can be positive or negative depending on the direction of change of attenuation, as  $x_A$  is positive for steatosis but negative for fibrosis and PDAC metastases. Thus, the sign enables better discrimination between steatosis and the other two diseases. Therefore, the  $C$  was used for a SVM training to classify liver states.

## 2.5 | SVM learning

The SVM<sup>27-29</sup> is a popular machine learning method related to the maximum margin classifier, which we employed in this study due to advantageous properties. To maximize the margins between distinct classes, the SVM considers those data points located close to class boundaries; these are called support vectors. This is an advantage with a small training data set, since the SVM does not require many data points distant from the class boundaries. Accordingly, we trained two SVM classifiers using 32 animals with 4 liver states: normal, PDAC tumor, fibrosis, and steatosis. All 2778 ultrasound scans in this study were analyzed for values (% blue,  $\alpha$ ,  $I_{dB}$ ,  $C$ ), and the features were averaged over frames for the same animal at the same time point,

resulting in 198 feature sets for all time points. However, the SVM training used only 64 feature sets which were obtained only from the first and last scans, which are called start and end time points, respectively. The start points were considered as normal livers ( $n = 32$ ), and the end points became disease groups, including steatosis ( $n = 12$ ), fibrosis ( $n = 11$ ), and PDAC tumor ( $n = 9$ ), which were assigned as 4 classes.

SVM implementation was performed using MATLAB (The MathWorks, Inc., Natick, MA, USA). For the classification of the four classes, multiple vs. one method was used, and the Gaussian kernel function was also used to construct smooth hyperplanes. Hence the SVM learning involves the two parameters of box constraint and sigma of the Gaussian, which were optimized to obtain robust hyperplanes, but avoiding overfitting; for the optimization, both hyperplane shapes and classification accuracies were observed by varying the two parameters.<sup>16</sup>

Two SVM trainings were performed with 1D ( $C$ ) and 3D (% blue,  $\alpha$ ,  $I_{dB}$ ) input features to suggest decision boundaries for trajectories based on parameters and images, respectively. When designing a liver diagnosis guide for clinicians with parameters, the simplest way can be providing decision boundaries for the 1D parameter which can be obtained repeatedly over time as the use of biparietal diameter.<sup>3</sup> Therefore, with the 1D combined parameter, we first trained a 1D SVM classifier producing 1D decision boundaries for liver states.

Moreover, when we try to consider higher dimensional parameters with more information,  $M$ -dimensional measurements can also be treated by PCA and SVM. Thus, we suggest a color imaging method with a 3D SVM classifier to effectively reveal and visualize disease trajectories with their decision boundaries; it is also applicable to  $M > 3$ . The 3D SVM training requires data normalization to avoid weights for specific features caused by data scales. The scales of each component (% blue,  $\alpha$ ,  $I_{dB}$ ) differ in ranges of [0, 100], [0, 1], [-60, 0], respectively, but % blue and  $\alpha$  hardly reach the extremes. (% blue,  $\alpha \times 100$ ,  $I_{dB}$ ) is likely to have comparable scales, which were therefore used for the 3D SVM training input and then produced 3D hyperplanes.

## 2.6 | Disease-specific imaging utilizing inner product and SVM

We propose *disease-specific imaging* to incorporate higher-dimensional features and show their changes using gradual colors, which can effectively visualize 3D feature trajectories caused by disease progressions. A framework for *disease-specific imaging* is illustrated in Figure 2(b, c), comprising the following steps: (a) Ultrasound data acquisition for a liver; (b) Feature measurements (% blue,  $\alpha$ ,  $I_{dB}$ ); (c) SVM classification for the input liver; (d) Color processing utilizing the SVM

and inner product; (e) Display of disease-specific images, identifying the liver state.

When training the SVM, we also performed a pre-processing for the inner-product-based color processing to obtain pathways of disease progressions in 3D space. The SVM trainings used 64 data at start and end time points, whereas the inner product operation employed features at all time points, consisting of 198 data of ( $\% \textit{blue}$ ,  $\alpha$ ,  $I_{dB}$ ). Using the 3D features, we define a vector  $v$  suggesting disease progression directions starting from a position  $N$  of normal liver:

$$v = (H, A, B) = (\% \textit{blue}, \alpha, I_{dB}) - N, \quad (6)$$

where  $N = (50\%, 0.5 \text{ dB/MHz/cm}, -15 \text{ dB})$ , and  $(H, A, B)$  represents feature changes from normal. When monitoring the vector  $v$  of a liver from normal state until any disease progresses,  $v$  starts from the origin of  $O = v = (0, 0, 0)$  and repeated measurements over time move gradually further from the origin  $O$ . Then, the trajectories of  $v$  indicate disease progression pathways and also directions of progression. By collecting  $v$  for each disease, linear pathways of tumor, fibrosis, and steatosis can be obtained by linear fittings; the lines are shown in Figure 3(a). Furthermore, the progression directions for tumor, fibrosis, and steatosis can be also estimated by calculating unit vectors of  $\hat{e}_1$ ,  $\hat{e}_2$ , and  $\hat{e}_3$ , respectively. Then, an inner product given by  $v \cdot \hat{e}_n$  indicates progression of  $n$ -th disease, and the values  $v \cdot \hat{e}_n$  for all enrolled 198 data are exhibited in Figure 3(a) as points with colors representing disease stages: the darker and brighter colors represent early and late stages of each disease, respectively, until the time of death given by the symbol “X”. Among the 198 projections onto the linear pathways, maximum values of the diseases were obtained:  $\textit{maxTumor} = \max(v \cdot \hat{e}_1)$ ,  $\textit{maxFib} = \max(v \cdot \hat{e}_2)$ , and  $\textit{maxFat} = \max(v \cdot \hat{e}_3)$  for tumor, fibrosis, and steatosis, respectively. The maximum values normalize the inner product results, and then as described in Figure 2(c), the normalized values of  $\frac{\langle v, \hat{e}_n \rangle}{\max(\textit{Disease})_n}$  are assigned for RGB color intensity for each pixel, wherein  $v$  is a measurement from any pixels within the ROI and the disease index of  $n$  is determined by the trained 3D SVM classifier. The normalized value utilizing inner product and SVM indicates disease progression; the higher the value, the more severe the disease, but values near 0 indicate normal or early stages. In this way, a color image is processed, and then overlaid with its B-scan; pixels classified as normal only display B-scan without color overlay.

### 3 | RESULTS

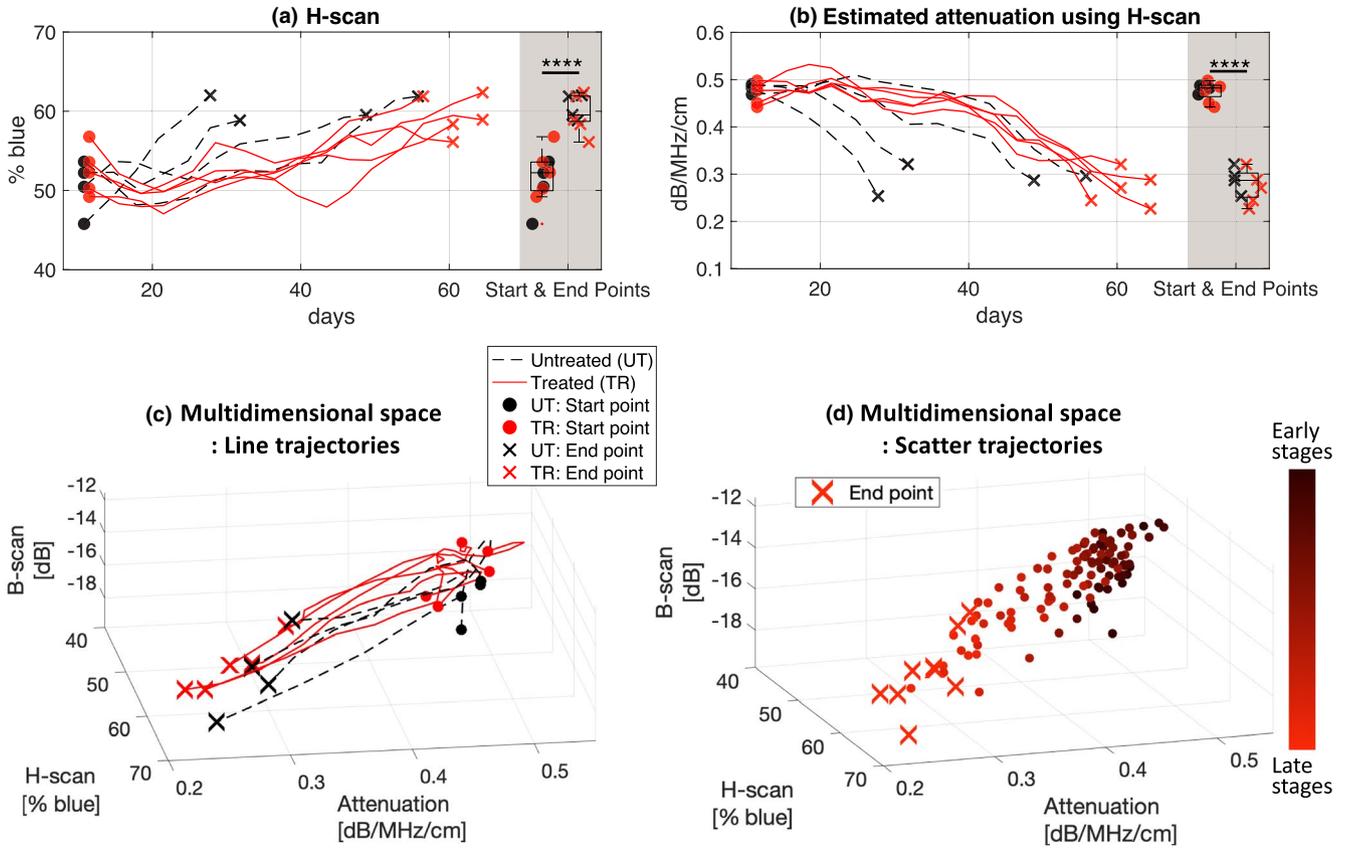
We present results obtained from three different animal models of disease, each study utilizing its own ultrasound scanner acquiring *in vivo* data. These are:

a liver fibrosis model, a liver steatosis model, and a PDAC metastatic model. The disease trajectories are based on H-scan analysis with three output parameters ( $\% \textit{blue}$ ,  $\alpha$ ,  $I_{dB}$ ) related to the upshift in scattering, the acoustic attenuation, and the B-scan echo intensity from regions within the liver, as described in the Methods section.

#### 3.1 | Trajectories in low dimensional spaces

Perhaps the most simple and traditional approach is measuring and plotting each parameter individually over time. For example, the progression of PDAC liver metastasis was found to produce changes in H-scan measures over time in days, indicating tumor growth as shown in Figure 4. Specifically, the  $\% \textit{blue}$  parameter in Figure 4(a) shows a gradual increase over time for both the untreated and treated (with chemotherapy) mice, but the chemotherapy group tends to have an increase in survival compared to the untreated group. When considering only start and end points, the percent difference is significantly different with  $P$ -value of  $5.7 \times 10^{-6}$ . Moreover, we estimated the attenuation coefficient  $\alpha$  using the H-scan approach by measuring frequency downshift along with depth, and Figure 4(b) demonstrates that  $\alpha$  can decrease with gradual tumor growth over time. The start and end points show significant difference ( $P$ -value =  $3.6 \times 10^{-10}$ ) without any overlap between the measurements of the two groups. However, the difference between the start and end points investigated by traditionally used B-scan intensity has a  $P$ -value of 0.01, which is statistically significant, but less significant than that of H-scan  $\% \textit{blue}$  or  $\alpha$ . Thus, H-scan  $\% \textit{blue}$  and attenuation estimation more sensitively describe the changes induced by the tumor compared to B-scan. In Figure 4(c,d), we combine the three metrics to make a joint assessment of the three, and therefore considered three dimensional (3D) features for each liver and time point, ( $\% \textit{blue}$ ,  $\alpha$ ,  $I_{dB}$ ). The gradually changed red colors from dark to bright red in Figure 4(d) illustrate start-to-end time points, representing liver changes from normal to an end-stage metastatic condition causing death. The trajectories in the plot of Figure 4(c) suggest a common pathway of tumor growth. To be specific, scatterers with features ( $\% \textit{blue}$ ,  $\alpha$ ,  $I_{dB}$ ) located near front left can be considered as late-stage tumor.

Unfortunately, 3D data are best appreciated with the assistance of an interactive graphics tool and 3D displays. Furthermore, in many studies, we have more than three measured parameters, including first order statistics, shear wave elastography, and contrast parameters.<sup>16,19,20</sup> In these cases, we cannot directly visualize a higher multi-dimensional space, and so require a strategy for reducing these. From mathematics, we have a number of options,



**FIGURE 4** H-scan trajectories indicating the progression of PDAC tumor metastasis. H-scan analysis extracted two features within liver tissues, which are blue percentage (a) and attenuation coefficient (b) along with time, representing days after tumor injection. The feature measurements were performed from start points (•) to end points (x), representing early and late stage (near death), respectively. The blue percentage (a) and attenuation (b) gradually increase and decrease over time, respectively, as tumor progresses. The features show significant differences between the early and late stages, and “\*\*\*\*” denotes  $p$ -value  $<0.0001$ . (c) A joint assessment of the two H-scan parameters and B-scan echo intensity in 3D parameter space, suggesting a pathway of tumor growth. (d) Similar information as (c) but with intermediate time points included. The gradual red colors from dark to bright represent early to late-stage tumors, which also demonstrates a tumor progression pathway [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

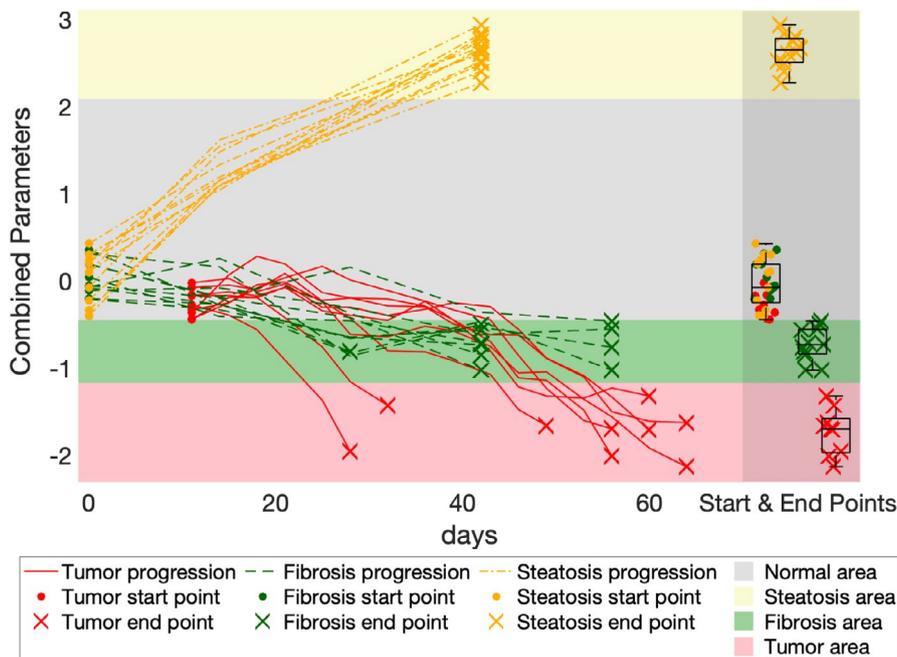
including simple projections from  $M$  dimensional space to 3D or 2D representations, or data-specific techniques such as principal component analyses.<sup>30,31</sup> These are linear operations; however, nonlinear transformations may be more revealing. To illustrate this effect, we show the three parameter measurements ( $\% \text{ blue}$ ,  $\alpha$ ,  $I_{dB}$ ) vs. time on the different classes of diseases from independent studies of liver fibrosis, steatosis, and PDAC metastases. To reduce the dimensionality, we applied a nonlinear transformation, providing combined parameter  $C$ , which was defined in the Methods section equation (5).

The nonlinear mapping creating the combined parameter  $C$  can be related to the distance in 2D from the center of the “normal” cluster, and creates a simplified interpretation of the trajectories that can be used clinically to assess the progress of any individual liver with any of these three conditions as shown in Figure 5. In this combined parameter plot vs. time, each disease has its own trajectory describing the changes of its tissue characteristics, and its end points formed clusters along the 1D axis of the combined parameter  $C$

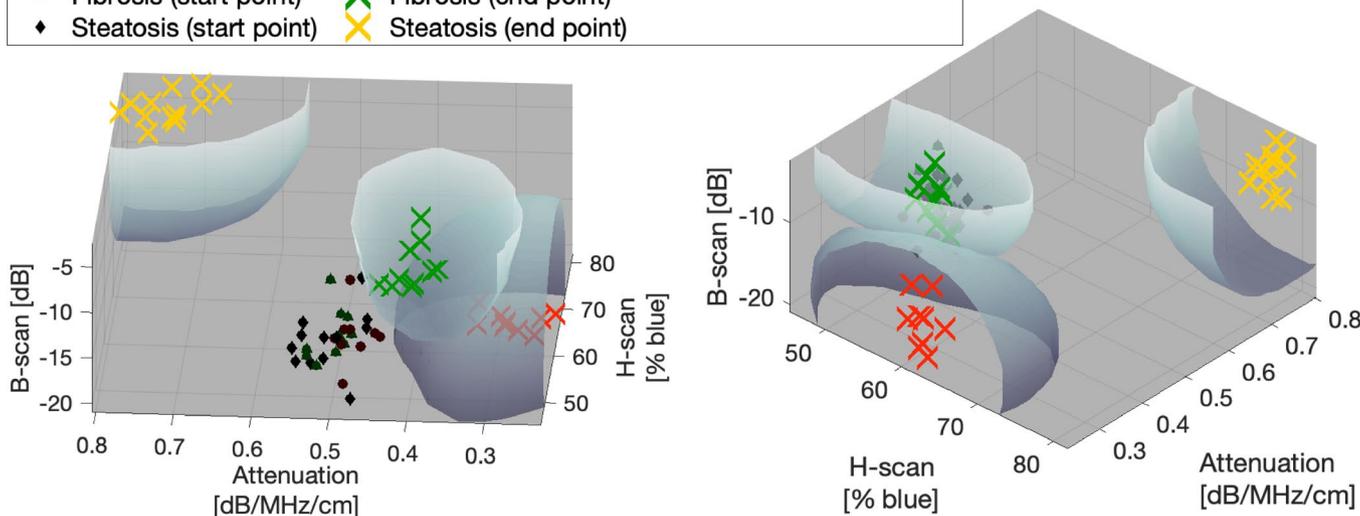
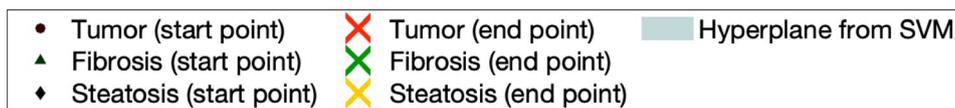
Accordingly, by using the end points of the three diseases in addition to the start points of all as a normal group, the four groups were used for 1D SVM training, producing 1D decision boundaries to differentiate the liver conditions: normal, fibrosis, steatosis, and PDAC metastases. The 1D classification accuracy is 100% without any overlap between the four groups. Hence, when a liver has a combined parameter  $C$  calculated from H-scan and attenuation estimation, clinicians may assess the progression of disease states based on mapping a particular patient's information over time.

### 3.2 | Clusters in multidimensional spaces for diagnosis and disease-specific imaging

Measurements in multidimensional space are provided in Figure 6, involving analysis of clusters. One effective means for this is the SVM.<sup>25,27-29</sup> Hyperplanes can be created by the SVM in the 3-parameter ( $\% \text{ blue}$ ,  $\alpha$ ,  $I_{dB}$ )



**FIGURE 5** Deriving a combined parameter  $C$  creates a simplified interpretation of the trajectories of the three diseases: fibrosis, steatosis, and PDAC metastases. Their trajectories were depicted using green, yellow, and red corresponding to fibrosis, steatosis, and PDAC metastases, respectively. By utilizing the start and end points of the combined parameters, 1D SVM provides the diagnostic classification based only on the combined parameters. Classification accuracy using the 1D SVM is 100% and the four groups are all statistically significant with  $p$ -value  $< 0.0001$  [Color figure can be viewed at wileyonlinelibrary.com]



**FIGURE 6** Hyperplanes constructed by SVM. The clusters of normal, fibrosis, steatosis, and PDAC metastasis were well separated with 100% classification accuracy. The original normal (start points) and endpoints of diseases are located within their hyperplane regions [Color figure can be viewed at wileyonlinelibrary.com]

space as shown in Figure 6. The SVM including the hyperplanes is also employed for disease-specific imaging.

These hyperplanes are able to discriminate the three studied conditions and normals with 100% accuracy. Furthermore, they define regions in  $M$ -dimensional spaces (for  $M$  measured parameters) associated with specific diseases; in this study,  $M = 3$ . The trajectory of

an individual with progressive disease, scanned repeatedly over time, may approach one of these clusters. When monitoring a liver growing any of the diseases, the trajectory starts from the normal cluster that has a center at 50% of H-scan, 0.5 dB/MHz/cm of attenuation coefficient, and a standardized  $-15$  dB of B-scan intensity, denoting (50, 0.5,  $-15$ ) in the 3D space. As the disease progresses, the trajectory reaches one of the hyperplane

boundaries defined by SVM, and then the trajectory goes further into the disease region, whereby the distance between the last measurement ( $\%blue$ ,  $\alpha$ ,  $I_{dB}$ ) and the center (50, 0.5, -15) of the normal cluster grows longer. The distance can be correlated with the vector  $(H, A, B) = (\%blue, \alpha, I_{dB}) - (50, 0.5, -15)$ .

For disease-specific imaging, this phenomenon can be visualized by using the inner product operation. First, linear paths from the normal center to the end points are provided in Figure 3(a); the three solid lines represent the central pathways of the diseases as defined by the clusters in Figure 6. With the  $(H, A, B)$ , we obtained the projections (a dot product in vector space) onto the paths, and the projections are shown as a scatter plot in Figure 3(a). The dark to bright colors correspond to progressively later stages. The earlier measurements were seen near the origin (0,0,0), and as the disease progresses, the assigned colors via dot product projection become more vivid. By employing the gradual increase in projection values caused by disease progressions, B-scan images can therefore be enhanced with the color intensity determined by the projection values. Examples of disease-specific imaging are displayed in Figure 3(b–d), (f–h), and (j–l). The top row shows fat growth with increase in yellow color. Similarly, the mid and bottom rows show fibrosis and tumor growth with increase in green and red colors, respectively. The diseases are progressive from left to right for the same animal, scanned repeatedly three times. Note that all the pixels at the earliest time point scans in Figure 3(b, f, j) do not have assigned colors, meaning that all the pixels were classified as normal without any disease. The given cases at mid time points in Figure 3(c, g, k), have some color-enhanced pixels, but these are less vivid than the colors at end time points in Figure 3(d, h, l). The final scans show nearly complete fill-in with saturated colors, meaning the pathologies were distributed throughout the entire ROIs. Figure 3(e, i, m) are histology images of the three disease models showing different liver compositions and patterns, which supports the three different colors for the disease models. In this framework, clinicians can assess disease based on the color enhancement, and moreover they can infer the severity and extent of the diseases based on color intensities. While scanning the liver repeatedly over time, they also can track the disease progression marked by gradual increase in color intensities.

## 4 | DISCUSSION

The H-scan analysis enables an assessment of disease progression by utilizing two approaches: defining multidimensional parameter trajectories that are disease-specific, and applying color-enhanced imaging representing the parameters' gradual changes. A simple

trajectory results from the combined parameter (C) derived from H-scan analyses and provides a useful tool for diagnosis or staging of the three conditions. H-scan trajectories more sensitively track the disease progression compared to other ultrasound imaging methods such as B-scan or shear wave elastography (SWE). Furthermore, H-scan yielded better performance than the recently used ultrasound application of SWE to detect early stage tumors.<sup>15</sup> Moreover, in addition to the high performance of H-scan, another important aspect for differential diagnosis is finding well-separated measurement clusters representing different liver conditions. Our combined parameter achieved 100% classification accuracy without any overlap according to SVM classification, thus it can provide a differential diagnosis between the three conditions of the liver that were included in this study. For visualization, the gradual changes of the H-scan parameters are converted to color intensities with reference to the multiparametric clusters and then the colors are transparently overlaid on B-scan images. The multidimensional parameters are efficiently unified into the color spaces by SVM and the inner product. Hence, the colors for diseases show up more vividly as the diseases become severe, potentially aiding clinicians in tracking disease progression.

This study enrolled three disease models and their livers were independently scanned by three ultrasound systems with different scan conditions such as different transducers, signal processing procedures, and parameter settings. The H-scan analysis investigated their ultrasound echoes and resulted in meaningful features, identifying liver tissues' characteristics of attenuation and structure provided by attenuation coefficient  $\alpha$  (dB/MHz/cm) and  $\%blue$ , respectively. There has been a long history of research in measuring  $\alpha$ , employing traditional<sup>32–34</sup> or currently developed<sup>35,36</sup> methods. Our H-scan analysis of attenuation within the three liver diseases of metastatic tumor, fibrosis, and steatosis showed consistent  $\alpha$  with the previous studies. For normal livers,  $\alpha = 0.5$  dB/MHz/cm; steatosis has a higher  $\alpha$  ( $>0.7$ ,<sup>33–36</sup> in this study  $\alpha = 0.74 \pm 0.04$  dB/MHz/cm), and fibrosis was reported to have  $\alpha < 0.45$ <sup>32,33</sup> (this study  $\alpha = 0.40 \pm 0.02$  dB/MHz/cm). Regarding the metastatic tumor model, our estimated attenuation ( $\alpha = 0.30 \pm 0.04$  dB/MHz/cm) is less than the normal and could be due to leaky tumor vessels and increase in fluid.<sup>37</sup>

We verified the disease-specific imaging by investigating the endpoints' histology figures (Figure 3(e, i, m)). Since the liver patterns in histology show widespread abnormal tissues in each case, a filling-in of colors within ROIs can be considered as correct results. Furthermore, the gradual increase or decrease in the measured parameters in Figures 4 and 5 was observed, whereby the gradual increase in color area and intensity in Figure 3 also appears correct. Particularly, Figure 3(c, g, k) have two types of regions

of B-scan-only and color-overlaid area, suggesting some normal regions interspersed with focal disease. However, we did not compare the same plane in the livers from ultrasound and histology images. Further studies for mid stage diseases to observe the identical planes of livers using corresponding ultrasound frames and histology sections are desired, which would verify the accuracy of our color imaging to detect disease area surrounded by normal tissues.

Our methodology including 1D diagnosis trajectory and disease-specific imaging was applied to the three H-scan outputs. However the approach can be extended to employ more parameters ( $M > 3$ ). For instance, there are methods producing additional parameters, such as SWE and histogram analysis. Let us assume we have  $M$  parameters generated by H-scan, B-scan, SWE, speckle histogram analysis, and possibly others, then PCA or nonlinear transformation can reduce the number of parameters, suggesting a combined 1D parameter for diagnosis. For disease-specific imaging, PCA can reduce the parameters into three principal components, and then the resulting methodology is the same as with this study. Alternatively, the  $M$  parameters can be directly used for imaging since SVM and inner product apply in a straightforward manner to  $M$ -dimensional inputs. Examples of multiparametric analysis ( $M > 3$ ) utilizing PCA and SVM are found in previous studies.<sup>16,19,20</sup>

One limitation of this study is that three distinct disease models were analyzed, however in clinical practice a patient can present with a combination of pathologies, for example fibrosis plus steatosis plus metastases plus other possible abnormalities. It remains to be seen how these combinations or regional differences can be resolved in terms of classification accuracy, spatial resolution, and color overlay fidelity with respect to gross pathology. Other changes in other soft tissues imaged with ultrasound also remain for future research.

To use the H-scan trajectory approach clinically in commercial ultrasound scanners, a simple protocol for clinicians and the possibility of real-time processing would be advantageous. Once users set an ROI box similar to the color Doppler scan, the rest of scanning procedure would be the same as a conventional B-scan. Additionally, when considering implementation on ultrasound devices, there is no requirement for additional or altered transmit or receive sequences, whereas Doppler modes or SWE require their own transmit, receive, and process sequences.

## 5 | CONCLUSIONS

To summarize, the H-scan analysis enables accurate quantification of ultrasound signals, and then its multiple parameters (along with any other available parameters such as shear wave speed or contrast-related measures) can be efficiently unified by parameter

transformation, SVM, and inner product. Consequently, H-scan trajectories assess the progression of pathologies using a single combined parameter over time, and through enhanced imaging with color overlays on B-scan derived from SVM clusters specific to disease states. These approaches resulted in 100% accuracy across different studies, classifying normal liver, PDAC metastasis, fibrosis, and steatosis within a unified framework. We anticipate clinical use of these methods since they achieved perfect discrimination between classes of disease and can be implemented in real-time.

## ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant R21EB025290. The authors are grateful to colleagues who conducted three different studies and who scanned liver models and provided RF data, including Rifat Ahmed and Marvin Doyley at the University of Rochester, Terri Swanson and Theresa Tuthill at Pfizer Inc., and Kenneth Hoyt and Lokesh Basavarajappa at the University of Texas at Dallas

## CONFLICT OF INTEREST

The authors have no relevant conflict of interest to disclose.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request, some data may be subject to third party restrictions.

## REFERENCES

1. Willlocks J, Donald I, Duggan TC, Day N. Foetal cephalometry by ultrasound. *J Obstet Gynaecol Br Commonw*. 1964;71:11-20.
2. Thompson HE, Holmes JH, Gottesfeld KR, Taylor ES. Fetal development as determined by ultrasonic pulse echo techniques. *Am J Obstet Gynecol*. 1965;92:44-52.
3. Häggström M. Medical gallery of Mikael Häggström 2014 (public domain). *WikiJ Med*. 2014;1. <https://doi.org/10.15347/wjm/12014.15008>.
4. Schreiber S, Dannhardt-Stieger V, Henkel D, et al. Quantifying disease progression in amyotrophic lateral sclerosis using peripheral nerve sonography. *Muscle Nerve*. 2016;54:391-397.
5. Hung OY, Molony D, Corban MT, et al. Comprehensive assessment of coronary plaque progression with advanced intravascular imaging, physiological measures, and wall shear stress: a pilot double-blinded randomized controlled clinical trial of nebivolol versus atenolol in nonobstructive coronary artery disease. *J Am Heart Assoc*. 2016;5:e002764.
6. Steffen BT, Guan W, Remaley AT, et al. Apolipoprotein B is associated with carotid atherosclerosis progression independent of individual cholesterol measures in a 9-year prospective study of Multi-Ethnic Study of Atherosclerosis participants. *J Clin Lipidol*. 2017;11:1181.e1181-1191.e1181.
7. Zaidman CM, Wu JS, Kapur K, et al. Quantitative muscle ultrasound detects disease progression in Duchenne muscular dystrophy. *Ann Neurol*. 2017;81:633-640.
8. Magistroni R, Corsi C, Martí T, Torra R. A review of the imaging techniques for measuring kidney and cyst volume in

- establishing autosomal dominant polycystic kidney disease progression. *Am J Nephrol*. 2018;48:67-78.
9. Dimcevski G, Kotopoulos S, Bjånes T, et al. A human clinical trial using ultrasound and microbubbles to enhance gemcitabine treatment of inoperable pancreatic cancer. *J Control Release*. 2016;243:172-181.
  10. Morales-Roselló J, Khalil A, Fornés-Ferrer V, et al. Progression of Doppler changes in early-onset small for gestational age fetuses. How frequent are the different progression sequences? *J Matern Fetal Neonatal Med*. 2018;31(8):1000-1008.
  11. Sacchi S, Dhutia NM, Shun-Shin MJ, et al. Doppler assessment of aortic stenosis: a 25-operator study demonstrating why reading the peak velocity is superior to velocity time integral. *Eur Heart J Cardiovasc Imaging*. 2018;19:1380-1389.
  12. Doris MK, Everett RJ, Shun-Shin M, Clavel MA, Dweck MR. The role of imaging in measuring disease progression and assessing novel therapies in aortic stenosis. *JACC Cardiovasc Imaging*. 2019;12:185-197.
  13. Sadeghi-Naini A, Sannachi L, Tadayyon H, et al. Chemotherapy-response monitoring of breast cancer patients using quantitative ultrasound-based intra-tumour heterogeneities. *Sci Rep*. 2017;7:10352.
  14. Li F, Huang Y, Wang J, et al. Early differentiating between the chemotherapy responders and nonresponders: preliminary results with ultrasonic spectrum analysis of the RF time series in preclinical breast cancer models. *Cancer Imaging*. 2019;19:61.
  15. Baek J, Ahmed R, Ye J, Gerber SA, Parker KJ, Doyley MM. H-scan, shear wave and bioluminescent assessment of the progression of pancreatic cancer metastases in the liver. *Ultrasound Med Biol*. 2020;46:3369-3378.
  16. Baek J, Poul SS, Swanson TA, Tuthill T, Parker KJ. Scattering signatures of normal versus abnormal livers with support vector machine classification. *Ultrasound Med Biol*. 2020;46:3379-3392.
  17. Quiaoit K, DiCenzo D, Fatima K, et al. Quantitative ultrasound radiomics for therapy response monitoring in patients with locally advanced breast cancer: multi-institutional study results. *PLoS One*. 2020;15:e0236182.
  18. Baek J, Swanson TA, Tuthill T & Parker KJ Support vector machine (SVM) based liver classification: fibrosis, steatosis, and inflammation. 2020 IEEE International Ultrasonics Symposium, 1-4, 2020. <https://doi.org/10.1109/IUS46767.2020.9251611>
  19. Basavarajappa L, Baek J, Reddy S, et al. Multiparametric ultrasound imaging for the assessment of normal versus steatotic livers. *Nat Sci Rep*. 2021;11:2655.
  20. Baek J, Poul SS, Basavarajappa L, et al. Clusters of ultrasound scattering parameters for the classification of steatotic and normal livers. *Ultrasound Med Biol* (in press). 2021. <https://doi.org/10.1016/j.ultrasmedbio.2021.06.010>
  21. Tufte ER. *The visual display of quantitative information*. Graphics Press; 1997.
  22. Mills BN, Connolly KA, Ye J, et al. Stereotactic body radiation and interleukin-12 combination therapy eradicates pancreatic tumors by repolarizing the immune microenvironment. *Cell Rep*. 2019;29:406-421.e405.
  23. Soares KC, Foley K, Olino K, et al. A preclinical murine model of hepatic metastases. *J Vis Exp*. 2014;51677.
  24. Dougherty G. *Digital image processing for medical applications, page 42*. Cambridge University Press; 2009.
  25. Parker KJ, Baek J. Fine-tuning the H-scan for discriminating changes in tissue scatterers. *Biomed Phys Eng Express*. 2020;6:045012.
  26. Jayalakshmi T, Santhakumaran A. Statistical normalization and back propagation for classification. *Int J Comput Theory Eng*. 2011;3:1793-8201. <https://doi.org/10.7763/IJCTE.2011.V3.288>.
  27. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273-297.
  28. Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Networks*. 1999;10:988-999.
  29. Bishop CM. *Pattern recognition and machine learning, Chapter 7*. Springer; 2006.
  30. Jolliffe IT. *Principal component analysis*, 2nd edn. Springer; 2002.
  31. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374:20150202.
  32. Kuc R. Clinical application of an ultrasound attenuation coefficient estimation technique for liver pathology characterization. *IEEE Trans Biomed Eng*. 1980;27:312-319.
  33. Garra SB, Shawker TH, Insana MF & Wagner RF. In vivo attenuation measurement: methods and clinical relevance. Proceedings of the Sixth EC workshop. Paris Ultrasonic Tissue Characterisation, p. 87-100, 1986.
  34. Parker KJ. Attenuation measurement uncertainties caused by speckle statistics. *J Acoust Soc Am*. 1986;80:727-734.
  35. Tada T, Iijima H, Kobayashi N, et al. Usefulness of attenuation imaging with an ultrasound scanner for the evaluation of hepatic steatosis. *Ultrasound Med Biol*. 2019;45:2679-2687.
  36. Gong P, Zhou C, Song P, et al. Ultrasound attenuation estimation in harmonic imaging for robust fatty liver detection. *Ultrasound Med Biol*. 2020;46:3080-3087.
  37. Fukumura D, Jain RK. Tumor microvasculature and microenvironment: targets for anti-angiogenesis and normalization. *Microvasc Res*. 2007;74:72-84.

**How to cite this article:** Baek J, Parker KJ. H-scan trajectories indicate the progression of specific diseases. *Med Phys*. 2021;48:5047–5058. <https://doi.org/10.1002/mp.15108>