Machine learning for automated hepatic fat quantification

H. Sagreiya*, A. Akhbardeh**, I. Durot,*** and D.L. Rubin****

*Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA, hersh.sagreiya@pennmedicine.upenn.edu
**Department of Radiology, Johns Hopkins University, Baltimote, MD, USA, akhbardeh@gmail.com
***Department of Radiology, Spital Emmental Hospital, Bungdorf, Switzerland, isabelle.durot@spital-emmental.ch
****Department of Radiology and Biomedical Data Science, Stanford, CA, USA rubin@stanford.edu

ABSTRACT

Nonalcoholic fatty liver disease is a global pandemic. This study investigated whether ultrasound point shear wave elastography measurements could predict hepatic fat quantification using a machine larning (ML) algorithm trained with fat quantification on MRI. 186 exams from Stanford and 50 from the University of Wisconsin were analyzed. Hepatic fat values were quantized into intervals of 5%, and a multi-model support vector machine (SVM) was run with 10 measurements of shear wave velocity as inputs. For each fat quantification level, a dedicated SVM was trained; the overall fat prediction was determined by fusing model results. Validation was via leave-one-out cross validation. Pearson correlation was calculated between predicted and actual fat quantification. There was a high correlation between ML-predicted fat quantification and MR-based fat quantification for Stanford (r=0.98) and Wisconsin (r=0.95) data. ML correctly predicted the fat quantification interval for most subjects.

Keywords: machine learning, ultrasound, MRI, hepatic fat, informatics

1 INTRODUCTION

Chronic liver disease due to nonalcoholic fatty liver disease (NAFLD) is a global health pandemic due to an increase in metabolic risk factors linked to the increasing prevalence of obesity worldwide [1][2], and 30% of the population is affected in the United States and Europe. It affects over one billion people worldwide and is projected to become the leading cause of liver transplantation [3][4]. Early diagnosis of NAFLD is crucial, as it can be treated before it eventually leads to hepatic inflammation, cirrhosis, liver failure, or hepatocellular carcinoma; in particular, it can be treated by addressing the underlying cause, whether through weight loss or insulin-sensitizing or antioxidant agents [5][6].

1.1 Hepatic Fat Quantification

Quantifying hepatic fat content was historically performed via histology, with an intracellular accumulation of fat droplets exceeding 5% of hepatocytes on liver biopsy being a gold standard for diagnosing hepatic steatosis [5][7][8]. The NASH Clinical Research Network scoring system for grading parenchymal invovement of steatosis includes the following categories: S0 (<5%), S1 (5-33%), S2 (34-66%), and S3 (>66%) [9]. However, biopsy is invasive and comes with its own set of risks, such as bleeding and infection, and there are limitations of biopsy such as sampling error. Hence, noninvasive methods to diagnose steatosis are warranted. Conventional ultrasound and computerized tomography are limited by low sensitivity and specificity, especially for mild steatosis, as well as the lack of a reliable quantification of steatosis [4][8]. Qualitative methods for assessing hepatic fat include hepatic echogenicity, ultrasound beam attenuation at greater depth, and poor visualization of either the venous structures or the diaphragm [10]. Magnetic resonance imaging measuring the proton density fat fraction (MRI-PDFF) has been shown in multiple studies to have a high correlation with histology-based steatosis grade and to have a high accuracy for quantifying steatosis, and it has been adopted as a reference standard for clinical trials; however, it is expensive and not widely available worldwide [4][11-15].

1.2 Ultrasound Elastography

Ultrasound elastography (USE) techniques such as point shear wave elastography (pSWE) and 2D shear wave elastography (2DSWE) are used in clinical practice for the management of patients with chronic liver disease. While USE has shown high accuracy for grading hepatic fibrosis [16], the effect of hepatic steatosis on USE measurements is unclear [17][18], and only a few studies have evaluated shear wave properties related to dispersion or viscoelastic models [18][19]. More studies are warranted to evaluate the potential for assessing hepatic steatosis by analyzing ultrasound properties such as shear wave elastography velocities. Machine learning, to date often applied to medical imaging data [20], may facilitate and improve this analysis, as it can potentially uncover patterns that are not readily perceived by human observers.

1.3 Study Purpose

The purpose of this study was to determine if shear wave velocity measurements using ultrasound point shear wave elastography can be used to predict hepatic fat quantification using a machine learning (ML) algorithm trained with fat quantification from MRI.

2 MATERIALS AND METHODS

This study was HIPPA-compliant and IRB-approved. Between March 2014 and July 2017, 186 examinations from 113 patients had results from both point shear wave elastography from a Siemens ultrasound scanner (Siemens Healthineers, Erlangen, Germany) and fat quantification from MRI. Fat quantification values were quantized into intervals of 5%, and a multi-model support vector machine (SVM) algorithm with the Gaussian radial basis function kernel was run with ten measurements of shear wave velocity as inputs. For each fat quantification interval, we trained a dedicated SVM model, and machine-learning based fat prediction was determined by fusing the results from all models. Intervals of 5% were chosen since the standard steatosis threshold is 5%. Results were validated using leave-one-out cross-validation. Next, for each quantization interval, the p-value for the fat prediction, as determined by the SVM, was calculated using a Wilcoxon rank-sum test. Finally, the correlation between predicted and actual fat quantification was done via Pearson correlation.

An additional similar such dataset was obtained from the University of Wisconsin for comparison, which contained 50 subjects, and a similar procedure was performed [5].

3 RESULTS

3.1 Stanford Dataset

For the Stanford dataset (**Figure 1**), there was high correlation between predicted fat quantification from ML and MRI-based fat quantification (r = 0.98). The dynamic range for MRI-based fat percent was between 0% and 35% for the Stanford dataset; thus, five SVM models had enough data samples to be trained. There was good score separation in intervals under 20% ($p < 10^{-9}$), whereas intervals greater than 20% had insufficient samples in this dataset (**Figure 2**). Finally, for each ultrasound examination (sample index), the predicted fat quantification interval was determined, and there was a high accuracy in predicting the correct fat quantification interval.

3.2 Wisconsin Dataset

As the Wisconsin dataset was smaller, it had a slightly lower Pearson correlation (r=0.95), once again with fewer data points at higher (>20%) fat quantification levels (**Figure 3**). Machine learning correctly predicted the fat quantification interval for most subjects.

A)

B)

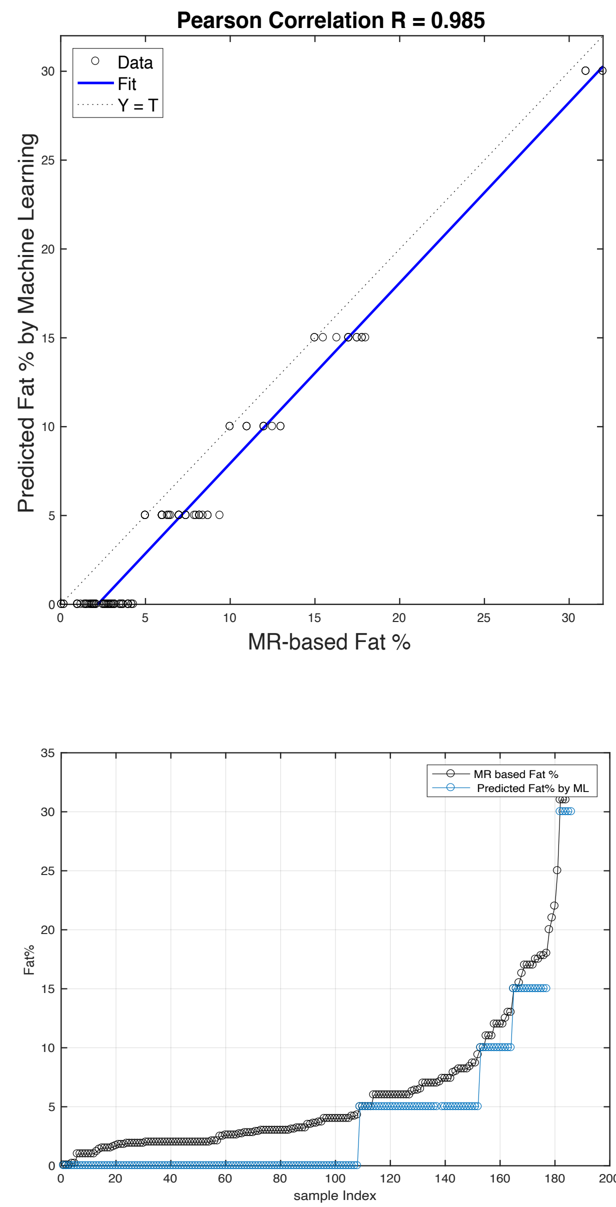
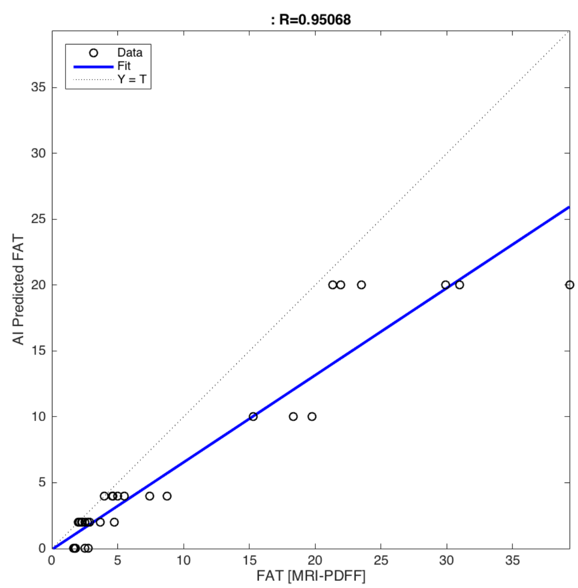


Figure 1: Stanford Dataset. A) Pearson correlation between MR-fat percent and the machine learning prediction. B) MR-fat percent (*black*) and predicted fat interval by machine learning (*blue*).

Interval	p-value	N
0-5	$2.05e^{-38}$	108
5-10	$2.10e^{-10}$	44
10-15	$2.34e^{-7}$	12
15-20	$1.43e^{-10}$	13
20-25	I.D.	3
25-30	I.D.	1
30+	$1.52e^{-05}$	5

Figure 2: Stanford Dataset. Strength of predictions at each fat quantification interval using the Wilcoxon rank-sum test. I.D. = Insufficient Data.

A)



B)

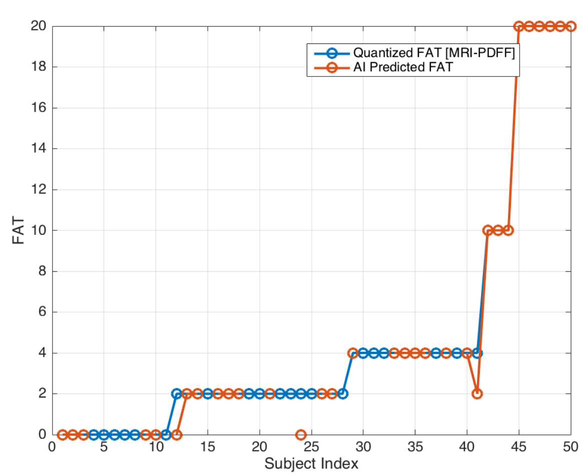


Figure 3: Wisconsin Dataset. A) Pearson correlation between MR-fat percent and the machine learning prediction. B) Comparing quantized MR-fat interval (*blue*) and machine-predicted fat interval (*red*).

4 DISCUSSION

Overall, this study showed that using data from point shear wave elastography in conjunction with machine learning was able to derive MRI-determined fat quantification levels. This was a notable result, as shear wave velocity has typically been used to predict hepatic stiffness rather than fat quantification. It is possible that there is informationn contained in the distribution of shear wave velocities that machine learning is able to leverage in order to predict hepatic fat content.

4.1 Prior Studies

Quantitative ultrasound has been studied for hepatic fat quantification in NAFLD, and parameters include hepatorenal index, speed of sound, attenuation coefficient, backscatter coefficient, controlled-attenuation parameter, and shear wave dispersion. One preliminary study of speed of sound in 17 patients resulted in an AUROC curve of 0.94 compared to MRI-PDFF and 0.95 compared to biopsy for classifying healthy versus fatty liver [21]. One 60-patient study employing attenuation coefficient found it to have a mean AUC of 0.79 for distinguishing between grade S1 versus S2+ steatosis [22]. This same 60-patient study found backscatter coefficient to have a mean AUC of 0.85 to distinguish between S1 versus S2+ steatosis. A meta-analysis of 19 studies found that the controlled-attenuation parameter (CAP) had an AUROC curve of 0.82 to distinguish between grade S1 versus S2+ steatosis [23]. However, CAP was not as accurate as MRI-PDFF [24][25]. Finally, shear wave dispersion showed an AUROC of only 0.47 [26]. Overall, while some of these quantitative ultrasound studies have been promising, many have involved small datasets or required non-standard ultrasound equiptment.

A few studies have explored using machine learning for ultrasound fat quantification. One study with 63 patients used deep learning to classify exams as either normal or fatty liver with 100% accuracy [27]. Another study with 55 patients utilized an Inception-ResNet-v2 network to classify exams as either normal or fatty liver, with an AUC of 0.98 [28]. Unlike those studies, this study sought to provide a hepatic fat quantification level, rather than simply classifying liver as normal versus abnormal. In addition, it was performed on datasets from two institutions.

4.2 Limitations and Future Directions

Regarding study limitations and future directions, one limitation was the sample size of this study. A future study on a larger cohort of patients from a more diverse array of institutions is warranted. In addition, this study specifically looked as shear wave velocity as the input to ML to predict MRI-determined fat quantification. A future study could make use of additional quantitative ultrasound parameters to investigate the extent to which they impact performance.

5 CONCLUSIONS

Machine learning, with shear wave velocity measurements from ultrasound point shear elastography as inputs, can be used to predict hepatic fat percent, which correlates well with MRI-determined fat quantification levels. Additional investigation and training with a bigger dataset are necessary to further validate the robustness of this approach.

6 ACKNOWLEDGEMENTS

We would like to thank Professor Scott Reeder for providing the Wisconsin dataset for comparison.

REFERENCES

[1] H. S. Leitao, S. Doblas, P. Garteiser, G. d'Assignies, V. Paradis, F. Mouri, C. F. Geraldes, M. Ronot and B. E. Van Beers, Radiology **283** (1), 98-107 (2017).
[2] S. C. Lin, E. Heba, T. Wolfson, B. Ang, A. Gamst, A. Han, J. W. Erdman, Jr., W. D. O'Brien, Jr., M. P. Andre, C. B. Sirlin and R. Loomba, Clin Gastroenterol Hepatol **13** (7), 1337-1345 e1336 (2015).
[3] R. Loomba, A. J. Sanyal, Nat Rev Gastroenterol Hepatol **10** (11), 686-690 (2013).
[4] M. R. Charlton, J. M. Burns, R. A. Pedersen, K. D. Watt, J. K. Heimbach, R. A. Dierkhising. Gastroenterology **141** (4), 1249-1253 (2011).
[5] H. Kramer, P. J. Pickhardt, M. A. Kliewer, D. Hernando, G. H. Chen, J. A. Zagzebski and S. B. Reeder, AJR Am J Roentgenol **208** (1), 92-100 (2017).
[6] S B. A. Neuschwander-Tetri, Clin Liver Dis **13** (4), 649-665 (2009).
[7] M. Blachier, H. Leleu, M. Peck-Radosavljevic, D. C. Valla and F. Roudot-Thoraval, J Hepatol **58** (3), 593-608 (2013).
[8] Z. Zhang, G. Wang, K. Kang, G. Wu and P. Wang, Sci Rep **6**, 32875 (2016).
[9] D. E. Kleiner, E. M. Brutt, M. V. Natta, C. Behling, M. J. Contos, O. W. Cummings, L. D. Ferrell, et al., Hepatology **41** (6), 1313-1321 (2005).
[10] A. E. Joseph, K. C. Dewbury, P. G. McGuire, The Br J Radiol **52** (615), 184-188 (1979).
[11] C. Caussy, S. B. Reeder, C. B. Sirlin and R. Loomba, Hepatology **68** (2), 763-772 (2018).
[12] I. S. Idilman, H. Aniktar, R. Idilman, G. Kabacam, B. Savas, A. Elhan, A. Celik, K. Bahar and M. Karcaaltincaba, Radiology **267** (3), 767-775 (2013).
[13] A. Tang, J. Tan, M. Sun, G. Hamilton, M. Bydder, T. Wolfson, A. C. Gamst, M. Middleton, E. M. Brunt, R. Loomba, J. E. Lavine, J. B. Schwimmer and C. B. Sirlin, Radiology **267** (2), 422-431 (2013).
[14] M. Di Martino, K. Koryukova, M. Bezzi and C. Catalano, Children (Basel) **4** (8) (2017).
[15] M. Noureddin, J. Lam, M. R. Peterson, M. Middleton, G. Hamilton, T. A. Le, R. Bettencourt, C. Changchien, D. A. Brenner, C. Sirlin and R. Loomba, Hepatology **58** (6), 1930-1940 (2013).
[16] Y. Lurie, M. Webb, R. Cytter-Kuint, S. Shteingart and G. Z. Lederkremer, World J Gastroenterol **21** (41), 11567-11583 (2015).
[17] G. Ferraioli, C. Tinelli, B. Dal Bello, M. Zicchetti, G. Filice, C. Filice and G. Liver Fibrosis Study, Hepatology **56** (6), 2125-2133 (2012).
[18] K. J. Parker, A. Partin and D. J. Rubens, Ultrasound Med Biol **41** (5), 1481-1487 (2015).
[19] T. Deffieux, J. L. Gennisson, L. Bousquet, M. Corouge, S. Cosconea, D. Amroun, S. Tripon, B. Terris, V. Mallet, P. Sogni, M. Tanter and S. Pol, J Hepatol **62** (2), 317-324 (2015).
[20] B. J. Erickson, P. Korfiatis, Z. Akkus and T. L. Kline, Radiographics **37** (2), 505-515 (2017).
[21] M. Imbault, A. Faccinetto, B. Osmanski, A. Tissier, T. Deffieux, J. Gennisson et al., Phys Med Biol **62** (9), 3582-3598 (2017).
[22] J. S. Paige, G. S. Bernstein, E. Heba, E. A. C. Costa, M. Fereirra, T. Wolfson, A. C. Gamst, et al., AJR Am J Roentgenol **208** (5), W168-W177 (2017).
[23] T. Karlas, D. Petroff, M. Sasso, J. Fan, Y. Mi, V. de Ledinghen, J Hepatol **66** (5), 1022-1030 (2017).
[24] K. Imajo, T. Kessoku, Y. Honda, W. Tomeno, Y. Ogawa, H. Mawatari, Gastroenterology **150** (3), 626-637.
[25] C. C. Park, P. Nguyen, C. Hernandez, R. Bettencourt, K. Ramirez, L. Fortney, et al, Gastroenterology **152**(3), 598-607.
[26] C. T. Barry, B. Mills, Z. Hah, R. A. Mooney, C. K. Ryan, D. J. Rubens, and K. J. Parker, Ultrasound Med Biol **38** (2), 175-182 (2011).
[27] M. Biswas, V. Kuppili, D. R. Edla, H. S. Suri, L. Saba, R. T. Marinhoe, J. M. Sanches, and J. S. Suri, Comput Methods Programs Biomed **155**, 165-177 (2017).
[28] M. Byra, G. Styczynski, C. Szmigielski, P. Kalinowski, L. Michalowski, R. Paluszkiewicz, et al., Int J Comput Assist Radiol Surg **13** (12), 1895-1903 (2018).