A Pilot Study for the Prediction of Liver Function Related Scores Using Breath Biomarkers and Machine Learning

Rakesh Kumar Patnaik

National Tsing Hua University

Yu-Chen Lin

National Tsing Hua University

Ashish Agarwal

National Tsing Hua University

Ming Chih Ho

National Taiwan University Hospital

J. Andrew Yeh (✉ jayeh@mx.nthu.edu.tw)

National Tsing Hua University

Research Article

Keywords: CTP, MELD, DMS, parameters

Posted Date: August 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-795823/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on February 7th, 2022.

See the published version at <https://doi.org/10.1038/s41598-022-05808-5>.

A PILOT STUDY FOR THE PREDICTION OF LIVER FUNCTION RELATED SCORES USING BREATH BIOMARKERS AND MACHINE LEARNING

Rakesh Kumar Patnaik¹, Yu-Chen Lin¹, Ashish Agarwal¹, Ming Chih Ho^{2,*}, and J Andrew Yeh^{1,*}

¹National Tsing Hua University, Institute of NanoEngineering and MicroSystems, Hsinchu, 30013, Taiwan

²National Taiwan University Hospital, Department of surgery, Taipei City 100225, Taiwan

*jayeh@mx.nthu.edu.tw

*mcho1215@ntu.edu.tw

Abstract

Liver function test is the first step to diagnose various liver diseases by measuring certain proteins and liver enzymes from the blood sample. After getting the required data of the clinical parameters from the blood test it is possible to calculate Child-Pugh (CTP), AST to PLT ratio (APRI) and Model for end-stage liver disease (MELD) clinical scores that help the doctors about the severity of the disease progression. Volatile organic compounds (VOCs) found in-breath and monitoring their concentration may be a prevailing method for disease diagnosis. In this work, Isoprene, Limonene, and Dimethyl sulphide (DMS) are considered as a potential breath biomarker related to liver disease. A dataset is designed, that includes the biomarkers concentration analysed from the breath sample before and after study subjects performed an exercise. Four regression methods are performed to predict the clinical scores using breath biomarkers data as features set by the machine learning techniques. Regression methods on the dataset for prediction are evaluated by mean absolute error and root mean square error. A significant difference observed for isoprene concentration ($p < 0.01$) and for DMS concentration ($p < 0.0001$) between liver patients and healthy subjects breath sample. Ensemble regression methods are found best suited for the dataset. The mean absolute error for CTP score is 0.07, for APRI score 0.1 and for MELD score 0.7. The R-square value between actual clinical score and predicted clinical score is found to be 0.78, 0.95, and 0.85 for CTP score, APRI score, and MELD score, respectively. These results suggest that breath biomarkers hold a promising approach for non-invasive test and mass screening related to liver disease.

Introduction

The liver is the body's prime organ and responsible for carrying out metabolism and protein absorption. The liver suffers from many kinds of diseases like hepatitis, fibrosis and cirrhosis. Liver cirrhosis and liver cancer stands as 11th and 16th rank as the most common cause of death globally with more than 1 million and 0.7 million deaths, respectively¹.

In comparison to the blood test, a breath test facilitates a non-invasive measurement where the concentration of target biomarkers reflects the change in body metabolism condition caused by sickness or infection. Breath VOCs concentration profile holds the information about body metabolism that is unique because VOCs in the blood stream transferred to the exhaled air and comes out. Hence, breath empowers the entire body blood inspecting. Any sort of inflammation, infection, presence of tumour cells and oxidative stress cause a change in the concentration of certain compounds found in the exhaled breath²⁻⁴.

In this research work, we target three VOCs, isoprene, DMS, and limonene. Various literature supports these three organic compounds may represent the change of liver function. In an in-vitro study it has been found that the isoprene is an intermediate outcome in the cholesterol synthesis mevalonate pathway⁵. The level of cholesterol in a healthy body is maintained in balance depending on food intake, cholesterol absorption and synthesis. Cholesterol biosynthesis is a complicated process that begins with acetyl-CoA enzyme and the first step is the

production of mevalonate from acetate. In further steps decarboxylation of mevalonate produces isoprene units which combines to form Squalene which finally converts to Cholesterol⁶⁻⁸. The severity of the liver disease builds a hepatic accumulation, increases HMG CoA reductase (HMGCR) which results in dysregulation of cholesterol metabolism⁹⁻¹². A diseased-liver function compare to a healthy-liver function may have an altered regulation of the said metabolism and we hypothesized that there will be some markable changes in isoprene concentration.

Isoprene concentration changes in human breath after exercise. In a study it has been found that isoprene concentration is more when heartrate increases after workout¹³. Adenosine triphosphate (ATP) is the primary high energy phosphate particle that empowers muscle compression. In the ATP production process, this acetyl-CoA enters the Krebs cycle to drive oxidative phosphorylation. This procedure, mainly happens in mitochondria, is the major source of ATP. In this procedure acetyl-CoA is produced which triggers the mevalonate cycle¹⁴. From previous studies as described, the concentration window of isoprene is broad^{15,16} and concentration of isoprene decreases as liver fibrosis stage increases towards critical condition¹⁷. In this work, instead of considering the absolute value, the change in isoprene concentration is recorded as described in the method section. The objective of the study is to observe the trend change in the isoprene concentration of healthy individuals and liver patients. For this reason, when the dataset is designed with the absolute values of isoprene concentration and the difference in isoprene concentration obtained before and after physical activity also included.

Liver is responsible for producing P450 metabolic enzyme. A cirrhotic liver got deprived of enough CYP2C19 compounds belongs to P450 enzyme that reduce the digestion of limonene and cause an increase in limonene levels in the breath¹⁸. Limonene is common to find in healthy person and liver patient breath because it is commonly found in food and beverages which gives citrus fruit flavour, and this is hard to avoid in regular diet¹⁹. Limonene is obvious to appear in the exhaled breath of both healthy and the patients' group, but the concentration remains higher in latter group.

Another VOC that holds the potential to be a liver breath biomarker is DMS. Breath DMS level increases when a liver deteriorates and leads to cirrhosis condition²⁰⁻²². Liver regulates synthesis and metabolism of hydrogen sulphide (H₂S). Methylation of H₂S produce DMS²³. There is no clear explanation of altered regulation of H₂S metabolism. But studies show that hepatic dysfunction disturbs the process by effecting the enzymes responsible for H₂S absorption. Part of H₂S converted to DMS²⁴. There is a possibility that a liver with progressed stage of dysfunction may cause change in the concentration of exhaled sulphur compounds. DMS is found significantly different in concentration between healthy and liver cirrhosis patients²⁵. Liver patients exhaled breath carries a higher amount of DMS than healthy individuals.

This work proposes a supervised machine learning regression approach that predicts score which are commonly used by doctors for liver diagnosis. Four different regression algorithms are employed to observe the prediction on a designed dataset that comprises of said VOCs concentration and this work also include the trend change of isoprene concentration because it varies after performing physical activity as input or feature set. For every regression method the output parameter is the real score and the model gets trained from training portion of the dataset and the score will be predicted from the trained model and test input data. The predicted score will be compared with the real score of respective study subject and an error matrix will narrate the performance of the chosen regression method.

Method

Study subject selection and clinical parameters:

The breath test protocol and research method ethics were approved by the Institutional Review Board of NTUH, Taipei, Taiwan (REC: 201912138RINB). All research was performed according to the ethics guidelines and signed informed consent form obtained from all study participants. Study subjects need to fit into inclusion criteria and if they considered being a part of our study, they need to sign the informed consent form. Breath test was done in NTUH, Taiwan. In the inclusion criteria the allowed age limit is 20 years to 70 years. Study subjects need to agree to perform staircase walking for two minutes and their heart rate must reach a minimum of 100 bpm. All study subjects need to maintain a minimum 8 hours of fasting before the breath test. Study subjects need to avoid smoking and alcoholic drink for 24 hours before the breath test. Exclusion criteria is that healthy subjects must not have past or present liver ailment issues and lungs disease. Patients are advised to withdraw if they feel any discomfort while following the breath protocol. Seventeen liver patients were enrolled in this study at NTUH, Taipei, Taiwan. Total seventeen patients that includes eight Patients had hepatic cellular carcinoma issues, six

patients admitted for surgery, two patients joined for the treatment of cholangiocarcinoma and four out of the considered patient list came for transplant evaluation. Twenty-eight healthy volunteers recruited as a healthy group and a liver function test is conducted to make sure no present symptoms or indication of the ill liver. Eight different common protein and liver enzymes which are commonly monitored in liver patients admitted in NTUH, Taipei considered in our study. This includes albumin (ALB), total bilirubin (TBIL), Aspartate transaminase (AST), Alanine transaminase (ALT), platelet count (PLT), prothrombin time (PT), normalized prothrombin time (INR), and creatinine (CRE). Table 1 describes the mean and standard deviation of said liver related protein and enzymes. Other than this information about hepatic encephalopathy (HE), ascites (ASC), serum sodium (Na), dialysis information (DLYS) also retrieved. Three clinical scores formula and calculation method are shown in supplementary documentation (Section1: Clinical score calculation). The following table explains all the relevant information.

<<Insert table 1>>

Breath collection protocol:

Before beginning the breath collection one prepared questionnaire is answered by the study subjects. The questionnaire gives us information about their brushing teeth, fasting, and cigarette smoking. To avoid the collection of target VOCs which may pre-exist in the mouth, rinsing the mouth before collection is essential.

One litre gas sampling aluminium bag is used for breath collection. Commercial breath collection mask has not been used to avoid any kind of discomfort to the patients. All samples were collected in a new bag. Breath samples were collected in three different bags in different situations. For the first bag of breath collection, study subjects were asked to first slow down their breathing until their heart rate reach 60 to 80 bpm. After the heart rate remains stable for about one minute, subjects are asked to blow into the bag (bag 1). The second bag (bag 2) is collected immediately after the subject performs staircase walking for a minimum of two minutes and the heart rate shoots up at least 100 bpm. The third bag is collected after 15 minutes of rest and after observing the heartrate is stable between 60 to 80 bpm (bag 3). After the breath collected in the bag within one-hour sampling process starts. Bag 1 and bag 3 is collected at rest for that reason the heartrate is expected to be at normal reading. The exercise section is systemized using 2 minutes of staircase walking after heartrate increases to 100 bpm on a same facility.

Breath analysis:

Breath samples are analysed using a GCMS-Clarus 680 instrument (PerkinElmer®, USA). At first sample from the bags were sampled into the 'Carbotrap' desorption tubes using a pump (SKC®) and then an Automatic Thermal Desorption (ATD) unit (TurboMatrix 350 PerkinElmer®, USA) that desorb the sample from the tube into the GC column. ATD comes with an automatic sample tube handling unit. That selects one tube at a time and heats the tube. Then sample transferred into a Trap unit then transferred to the GC column²⁶.

<<Insert figure 1>>

The VOCs concentration was calculated utilizing a GCMS attached with a programmed desorption unit. The ATD has a desorption unit and a facility for the build-up of low-concentrate VOCs. The GC/MS station utilized helium gas of 99.9995% pure as the transporter gas. The GC section was an Elite 624 polar column (PerkinElmer, USA) with length, inside breadth, and film thickness is 60 meters, 0.25 mm and 1.4 µm respectively. This GC has a single section. The temperature ramp-up of the profile has three stages. At first, it is 35-degree C for five minutes. At that point at a pace of 7-degree C /min, it expanded up to 90 degree C and afterward, the rate expanded to 30-degree C /min until it arrived at 220-degree C. The temperature needs to rise gradually so the peaks show up at some point and a moderate temperature ramp-up profile keeps up to focus for those mixes with less boiling point. Abrupt increment in column temperature cause short retention time and generally the peaks gets overlapped. The mass-to-charge proportion (m/z) was considered between the range of 20 and 300. VOCs were distinguished utilizing the mass information base designed by NIST provided by GC-MS software. Quantification is done by peak area comparison with reference to the calibration done using the standard compounds.

Results

Figure 2 shows the mean and standard deviation (SD) plot of isoprene concentration from three different bags from the healthy and patient group. As said earlier the range of isoprene is wider and overlap between the two groups. To observe the isoprene behaviour in the said groups the trend change in the concentration profile is the priority of this work.

<<Insert figure 2>><<Insert figure 3>>

The concentration of isoprene difference before and after a workout is shown in figure 3 and the difference in the concentration observed. The mean limonene concentration for all subjects in bag1, bag2 and bag3 found as 15 ppb, 14.92 ppb and 13.53 ppb. As per the designed breath collection protocol the concentration of limonene remains almost the same in the three bags, but the concentration of limonene is different in the patient group and the healthy group as shown in figure 4. Similarly, the mean DMS concentration in three bags are 1.0, 1.05 and 0.69 ppb but mean concentration of DMS in the patient group and the healthy group (figure 5) are 2.2 ppb and 0.2 ppb, respectively.

<<Insert figure 4>> <<Insert figure 5>>

Statistical analysis

An independent one-tailed t-test conducted on the target VOCs concentration data collected from both the groups and a dependent one-tailed t-test performed on the isoprene concentration before and after performing physical activity on the total study subjects. An independent t-test is performed to observe the existence of a significant difference between patients (n=17) and healthy group (n=28) with significance level = 0.01. The mean concentration of isoprene from each of the three bags are tested including the difference in isoprene concentration before and after performing a physical activity with mean limonene and mean DMS concentration between two groups. A dependent t-test is performed on entire population between bag1 and bag2 that indicates the possibility of significant difference in Isoprene concentration after physical activity. Similarly, another dependent t-test is performed on entire population between bag2 and bag3 that indicates the possibility of a significant difference in isoprene concentration after rest.

<<Insert table 2>>

Table 2 explains that the target VOCs shows different concentration profile in both the groups and isoprene compound which is a by-product of cholesterol synthesis shows a trend change in the concentration.

Regression analysis

Dataset:

Three different datasets are prepared as CTP score, APRI score and MELD score as output variable, respectively. There are seven breath parameters as input features kept constant in all the three datasets. Those features are isoprene concentration from three different bags, isoprene concentration difference from bag2 to bag1 and bag2 to bag 3, mean limonene concentration and mean DMS concentration. All features are numerical and carry no missing value. Distribution plot for the three clinical scores shown in supplementary figure 1-3.

Grid search and hyper-parameter tuning

Four different regression methods applied in this work, are Linear regression (LR), Support Vector Regression (SVR), Random Forest Regression (RFR) and Extra Tree Regression (ETR). These methods are implemented using 'Jupyter 6.0.1' version for Python code development. The simple and the classic method is linear regression is as a baseline regressor. Support Vector Machine which very popular in creating a hyperplane and forming a decision boundary. Ensembles learning method also applied to the dataset for score prediction.

Usually grid search is a time saving option for parameter tuning. In this work, grid search is run for SVR, RFR and ETR algorithms for three different datasets. The parameters and their optimum value with scoring (negative mean absolute error) shown in supplementary table 2. Using the grid search function, the optimum value for shown parameters of every said method for three different scores shown. The performance of regression analysis is evaluated based on mean absolute error (MAE) and root mean squared error (RMSE). The test size during APRI

score prediction kept 20% whereas for CTP score and MELD score prediction test size was 30%. The test size kept independent from the training data.

Usually the predicted score using different regression method gives a floating number. The nomenclature of medical scores (CTP and MELD) are whole number. APRI score is one digit after the decimal. For this reason, the predicted scores are rounded off to nearby whole number in the case of CTP and MELD. The predicted APRI score is rounded to one digit after the decimal.

Regression result

Table 3 shows the performance parameters of said regression analysis methods for CTP score prediction. The mean value of the actual CTP score is 5.5 and the minimum MAE is 0.071 (ETR method) which is 1.3 % of the mean CTP score.

<<Insert table 3>>

Table 4 shows the model scores of chosen regression analysis methods for APRI score prediction. The mean value of the actual APRI score is 0.5 and the MAE is 0.118 which is 23 % of the mean APRI score. The distribution of APRI score is more distributed as shown in the figure.

<<Insert table 4>>

Table 5 shows the error values of selected regression analysis methods for the MELD score prediction. The mean value of the actual MELD score is 7.5 and the MAE is 0.785 which is 10.4 % of the mean MELD score.

<<Insert table 5>>

Out of four methods, ETR method is most suitable to the dataset and the MAE is comparatively less for the CTP and the APRI score prediction. Whereas, RFR method is most suitable to the dataset and MAE is comparatively less for the MELD score prediction. Figure 6 scatter plot shown using the actual score and predicted CTP score using ETR method.

<<Insert figure 6>>

Figure 6 scatter plot shows the counts or number of test subjects (digit shown next to marker) for overlapped entries. As the dataset distribution shows CTP score 5 is more in number. All healthy person and patient with liver disease at minor stage their CTP score is 5 which the minimum value. The count number shows the entries in that co-ordinates. R-squared value is more than 0.78. As the CTP score states that 5 to 6 is type A and 7 to 9 is type B. Above plot shows that there is one misclassification where the actual score is 6 and predicted score is 5. Still for the CTP score 5 if it got misclassified to 7 a type A patient will be predicted as type B.

<<Insert figure 7>>

The test size while predicting APRI score kept 20% of the total dataset. The APRI score states that if the value less than 0.5 there is no fibrosis present and more than 1.5% strong chance of fibrosis. All points very well predicted with an R-squared value more than 0.95 as shown in figure 7.

<<Insert figure 8>>

The R-squared value is also more than 0.85 as shown in figure 8. The number of wrong predictions can be improved with an even distribution covering all possible score values. Figure 9 shows all the said regression method R-squared value for three different scores.

<<Insert figure 9>>

In this work ensemble regressors performed a better prediction. Both RFR and ETR did a good match with the actual score. This study shows there is a possibility of predicting clinical scores using supervised machine learning regression methods, where the predicted scores narrate liver condition or liver disease stage using breath biomarkers. With further increasing a dataset with a distribution that fits the real world this model will be reliable for score prediction.

Discussion

From every study subject three bags of breath sample collected, and three target VOCs are quantified from each sample bag. Isoprene concentration after performing the exercise is higher than the isoprene concentration from the sample collected before exercise¹³ and the sample collected after rest. The isoprene concentration profile of the healthy and patients group observed for a significant difference and as described in the literature with increase in fibrosis stage isoprene concentration decreases¹⁷. Limonene and DMS concentrations quantified from each sample remain almost similar in three bags for that reason mean concentration of limonene and DMS are considered in the dataset.

An independent t-test is performed on both groups, a significant difference ($p < 0.01$) is observed for isoprene, and DMS concentration, as shown in Table 2. This t-test result describes the chosen VOCs concentration is different in liver patients and healthy person group. An ill liver got deprived of an efficient metabolism potential mostly due to damage of active liver cells. Changes in metabolism alters the VOCs concentration. Figure 2 shows the isoprene concentrations profile from the three breath samples from both groups is different. The isoprene concentration at rest (bag 1) and after rest (bag 3) in the liver group is lower compare to the healthy group. After performing exercise collected sample (bag 2) isoprene concentration is about two times higher than bag 1 concentration but in the case of a liver patient, it is not that much higher. A dependent t-test performed on the isoprene concentration difference calculated from bag 1 to bag 2 and bag 2 to bag 3. This test describes the increase in isoprene concentration after performing physical activity and decreases in isoprene concentration after performing rest. A high significant value ($p < 0.0001$) observed in the test.

The mean concentration of limonene which usually does not increase after performing exercise remains nearly same in all the three bags, so the mean limonene concentration from the samples of both the group considered. The mean limonene concentration of the healthy group is less compared to the liver patient group as shown in figure 4 and support a previous study¹⁸. The same trend also observed with DMS concentration in the breath^{22,25}. DMS concentration in the breath is supposed to be higher in liver patients than a healthy people. DMS mean concentration found to be significantly different ($p < 0.001$) and higher also in this study (figure 5).

Three datasets designed, each comprises of seven input features and the predicted output variable for each dataset is CTP score, APRI score, and MELD score, respectively. Ensemble learning-based ETR did a better prediction of scores compared to the other three said regression methods. Two different errors were used to evaluate each regression method. As shown in table 3 mean value of CTP in the dataset is 5.5 and using ETR the least MAE and RMSE is found to be 0.071 and 0.267, respectively and which makes less than 5% of the mean CTP score. There is one misclassification and the r-squared value is 0.78 between real and predicted CTP score plotted in figure 6.

Table 4 shows the APRI score prediction the MAE and RMSE are 0.118 and 0.168, respectively. The mean APRI score is 0.5 and the errors are more than 10% of this value. APRI score is more distributed compared to other scores. The number of misclassifications in APRI score prediction is shown in the figure 7 but the terminology of APRI score says a safe limit less than 0.5 and a risk level starts from 1.5. The prediction well fitted in these limits. The r-squared value between the actual and predicted score is 0.95. There is a possibility with a greater number of samples covering the entire range may improve the prediction.

MELD data are very well distributed compared to CTP data. The MELD describes end-stage liver disease score and it dictates the mortality percentage. For MELD score the mean value is 7.5, the MAE and RMSE are found to be 0.785 and 1.101, respectively as shown in table 5. The error is less than 15% of the mean MELD value of the dataset. Three misclassifications observed where the real value is 10 and the predicted score is 8. Figure 8 shows the r-squared value of 0.85 between the real and predicted score.

Figure 9 put together the performance of four different regression methods for predicting the scores. The ensemble learning fits very suitable to this dataset and predicts score very close to the real scores.

Conclusion

This pilot study has shown a promising result with significant different breath profiles between liver patients and healthy volunteers. The use of machine learning for the prediction of scores is found very promising for use of breath biomarkers for liver function diagnosis. More than one biomarker and multiple methods employed for score

prediction made this study more reliable. These findings form a stepping stone towards non-invasive mass screening diagnosis of liver function using breath biomarkers. This study also shown the ability of machine learning for the prediction of liver ailment stage based on breath biomarkers data.

References

1. Asrani, S. K., Devarbhavi, H., Eaton, J. & Kamath, P. S. Burden of liver diseases in the world. *J. Hepatol.* **70**, 151–171 (2019).
2. Jansson, B. O. & Larsson, B. T. Analysis of organic compounds in human breath by gas chromatography-mass spectrometry. *J. Lab. Clin. Med.* **74**, 961–966 (1969).
3. Conkle, J. P., Camp, B. J. & Welch, B. E. Trace composition of human respiratory gas. *Arch. Environ. Health* **30**, 290–295 (1975).
4. Gelmont, D., Stein, R. A. & Mead, J. F. Isoprene - The main hydrocarbon in human breath. *Biochem. Biophys. Res. Commun.* **99**, 1456–1460 (1981).
5. Deneris, E. S., Stein, R. A. & Mead, J. F. In vitro biosynthesis of isoprene from mevalonate utilizing a rat liver cytosolic fraction. *Biochem. Biophys. Res. Commun.* **123**, 691–696 (1984).
6. Cerqueira, N. M. F. S. A. *et al.* Cholesterol Biosynthesis: A Mechanistic Overview. *Biochemistry* **55**, 5483–5506 (2016).
7. Liscum, L. Cholesterol biosynthesis. *Biochem. Lipids, Lipoproteins Membr.* 399–421 (2008) doi:10.1016/B978-044453219-0.50016-7.
8. Das, S., Pal, S. & Mitra, M. Significance of Exhaled Breath Test in Clinical Diagnosis: A Special Focus on the Detection of Diabetes Mellitus. *J. Med. Biol. Eng.* **36**, 605–624 (2016).
9. Puri, P. *et al.* A lipidomic analysis of nonalcoholic fatty liver disease. *Hepatology* **46**, 1081–1090 (2007).
10. Arvind, A., Osganian, S. A., Cohen, D. E. & Corey, K. E. Lipid and Lipoprotein Metabolism in Liver Disease. in (eds. Feingold, K. R. *et al.*) (2000).
11. Serviddio, G., Bellanti, F. & Vendemiale, G. Free radical biology for medicine: learning from nonalcoholic fatty liver disease. *Free Radic. Biol. Med.* **65**, 952–968 (2013).
12. Min, H.-K. *et al.* Increased hepatic synthesis and dysregulation of cholesterol metabolism is associated with the severity of nonalcoholic fatty liver disease. *Cell Metab.* **15**, 665–674 (2012).
13. Salerno-Kennedy, R. & Cashman, K. D. Potential applications of breath isoprene as a biomarker in modern medicine: A concise overview. *Wien. Klin. Wochenschr.* **117**, 180–186 (2005).
14. Gruenbacher, G. & Thurnher, M. Mevalonate metabolism in immuno-oncology. *Front. Immunol.* **8**, 1–8 (2017).
15. Fenske, J. D. & Paulson, S. E. Human breath emissions of VOCs. *J. Air Waste Manag. Assoc.* **49**, 594–598 (1999).
16. Alkhouri, N. *et al.* Analysis of breath volatile organic compounds as a noninvasive tool to diagnose nonalcoholic fatty liver disease in children. *Eur. J. Gastroenterol. Hepatol.* **26**, 82–87 (2014).
17. Alkhouri, N. *et al.* Isoprene in the exhaled breath is a novel biomarker for advanced fibrosis in patients with chronic liver disease: A pilot study. *Clin. Transl. Gastroenterol.* **6**, e112-7 (2015).
18. Fernández del Río, R. *et al.* Volatile Biomarkers in Breath Associated With Liver Cirrhosis - Comparisons of Pre- and Post-liver Transplant Breath Samples. *EBioMedicine* **2**, 1243–1250 (2015).
19. Friedman, M. I. *et al.* Limonene in expired lung air of patients with liver disease. *Dig. Dis. Sci.* **39**, 1672–1676 (1994).
20. Kolluru, G. K., Shen, X., Bir, S. C. & Kevil, C. G. Hydrogen sulfide chemical biology: pathophysiological roles and detection. *Nitric oxide Biol. Chem.* **35**, 5–20 (2013).
21. Wu, D.-D. *et al.* Hydrogen Sulfide as a Novel Regulatory Factor in Liver Health and Disease. *Oxid. Med. Cell. Longev.* **2019**, 3831713 (2019).
22. Tangerman, A., Meuwese-Arends, M. T. & Jansen, J. B. Cause and composition of foetor hepaticus. *Lancet (London, England)* vol. 343 483 (1994).
23. Łowicka, E. & Bętkowski, J. Hydrogen sulfide (H₂S) - The third gas of interest for pharmacologists. *Pharmacol. Reports* **59**, 4–24 (2007).
24. Campbell, A. W. Nitric oxide. *Adv. Mind. Body. Med.* **27**, 4–5 (2013).
25. Van den Velde, S., Nevens, F., Van hee, P., van Steenberghe, D. & Quirynen, M. GC-MS analysis of breath odor compounds in liver patients. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **875**, 344–348 (2008).
26. Arnts, R. R., Singsaas, E. L. & Sharkey, T. D. Gas Chromatographic Techniques for the Measurement of Isoprene in Air. **1966**, 47–71 (1997).

Acknowledgements

We thank members of Dr. Ming Chih Ho laboratory for helping us to carry out the test and patient recruitment. We also like to acknowledge Dr. Chen kuo Chiang for his valuable suggestion about data analysis. This work supported from Ministry of Science and Technology, Taiwan research project funding (Project Grant: MOST 107-2221-E-007-110-MY3 and MOST 110-2221-E-007-068-MY3). All study subjects blood test conducted at NTU Hospital, Taiwan. The Institutional Review Board passed this study design, reference REC: 201912138RINB.

Author contributions statement

Rakesh Kumar Patnaik did the study design, test the samples and analysed the data. Yu-Chen Lin and Ashish Agarwal help to do equipment setup and sample collection. Ming Chih Ho verified the study design, look after the data collection and confirm the health status of study participants. J. Andrew Yeh conceived, guided and analyzed the results. All authors reviewed the manuscript.

Additional Information

The authors declare no competing interest.

Table 1. Demographic and Clinical parameters of study subjects, this contains mean and standard deviation of shown parameters. Data shown for patients and healthy group.

	Healthy (n=28)	Liver patient (n=17)	Total (45)
Age	24.78±5.07	59.76±7.31	38±18.15
Height	168.03±8.75	163.02±7.6	166.14±8.6
Weight	63.85±9.87	69.18±13.14	65.86±11.38
Albumin	4.64±0.36	4.1±0.69	4.4±0.57
T-bilirubin	0.74±0.29	1.25±0.95	0.93±0.66
AST	18.85±4.64	46.23±20.04	29.2±18.42
ALT	15.64±9.17	44.23±28.73	26.44±23.41
Platelet count	252.17±57.51	156.41±97.83	216±87.83
Prothrombin time (PT)	10.91±0.48	11.3±0.81	11.06±0.64
Normalized PT	1.05±0.04	1.07±0.07	1.06±0.06
Creatinine	0.84±0.17	0.81±0.15	0.83±0.16

Table 2. Breath test parameters and independent t-test result

Type of test	Parameter	t value	p value
Independent t-test on patients (n=17) and healthy group (n=28)	Isoprene_bag1	2.691	0.005
	Isoprene_bag2	3.042	0.002
	Isoprene_bag3	3.287	0.001
	Mean limonene	-1.165	0.125
	Mean DMS	-6.857	0.00001
Dependent t-test on total study subjects (n=45)	Isoprene concentration before and after workout (bag1 and bag2)	9.239	0.00001
	Isoprene concentration after workout and rest (bag2 and bag3)	-9.238	0.00001

Table 3. Performance of different regression methods for prediction CTP score

	MAE	RMSE
LR_CTP	0.362	0.557
SVR_CTP	0.285	0.654
RFR_CTP	0.285	0.534
ETR_CTP	0.071	0.267

Table 4. Performance of different regression methods for prediction APRI score

	MAE	RMSE
LR_APRI	0.406	0.479
SVR_APRI	0.248	0.313
RFR_APRI	0.271	0.405
ETR_APRI	0.118	0.168

Table 5. Performance of different regression methods for prediction MELD score

	MAE	RMSE
LR_MELD	1.341	1.626
SVR_MELD	1.071	1.488
RFR_MELD	0.857	1.195
ETR_MELD	0.785	1.101

Figure 1. Layout of breath collection and analysis setup

Figure 2. Isoprene concentration (ppb) from collected exhaled breath of patient and healthy group (Bag1: before physical activity, bag2: After physical activity, bag3: after 15 minutes of rest)

Figure 3. Isoprene concentration (ppb) difference from collected exhaled breath of patient and healthy group (Bag2-1: before physical activity and after physical activity, bag2-3: after physical activity and after 15 minutes of rest)

Figure 4. Mean limonene concentration (ppb) difference from collected exhaled breath of patient and healthy group

Figure 5. Mean DMS concentration (ppb) difference from collected exhaled breath of patient and healthy group

Figure 6. Comparison of actual and predicted CTP score (using ETR) and number of instances shown as label

Figure 7. Comparison of actual and predicted APRI score (using ETR) and number of instances shown as label

Figure 8. Comparison of actual and predicted MELD score (using ETR) and number of instances shown as label

Figure 9. R-squared value of actual score to predicted score using four different regression method for CTP, APRI and MELD score.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryinformation.pdf](#)