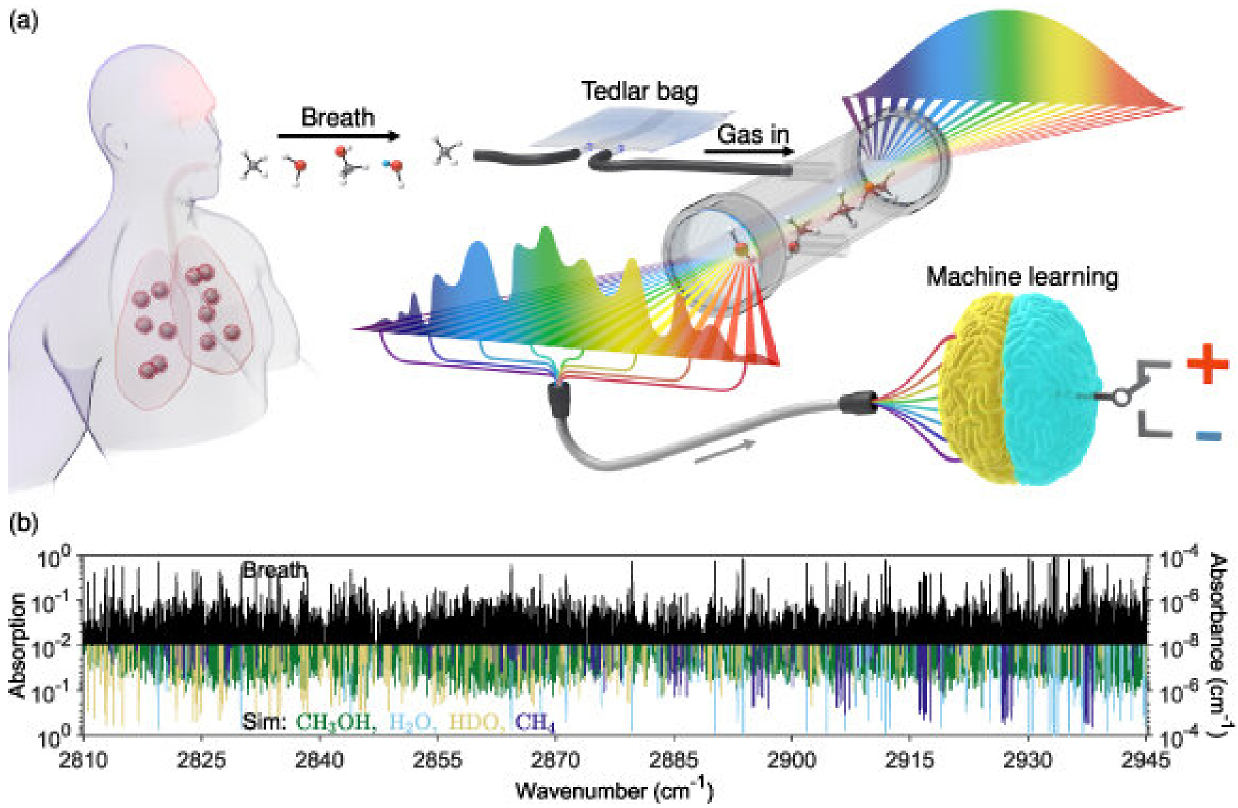
PHYS.ORG

New frequency comb breathalyzer detects COVID-19 with excellent accuracy

April 6 2023



CE-DFCS breathalyzer. (a) Schematic representation of the working principle of the device. An exhaled human breath sample was collected in a Tedlar bag and then loaded into an analysis chamber. The chamber was surrounded by a pair of high-reflectivity optical mirrors. A mid-infrared frequency comb laser interacted with the loaded sample and generated a broadband molecular absorption spectrum. The spectroscopy data was then used for supervised machine learning analysis to predict the binary response class for the research subject (either positive or negative). (b) Sample absorption spectrum collected from a research subject's exhaled breath (black). Inverted in sign and plotted with different colors are four fitted species (CH_3OH, H_2O, HDO, and CH_4) that give the most dominant absorption features. Credit: *Journal of Breath Research* (2023). DOI: 10.1088/1752-7163/acc6e4

JILA researchers have upgraded a breathalyzer based on Nobel Prize-winning frequency-comb technology and combined it with machine learning to detect SARS-CoV-2 infection with excellent accuracy in 170 volunteer subjects. Their achievement represents the first real-world test of the technology's capability to diagnose disease in exhaled human breath.

Their study on this topic was published in the *Journal of Breath Research*.

Frequency comb technology has the potential to non-invasively diagnose more health conditions than other breath analysis techniques, while also being faster and potentially more accurate than some other medical tests. Frequency combs act as rulers for precisely measuring different colors of light, including the infrared light absorbed by biomolecules in a person's breath.

Human breath contains more than 1,000 different trace molecules, many of which are correlated with specific health conditions. JILA's frequency comb breathalyzer identifies chemical signatures of molecules based on exact colors and amounts of infrared light absorbed by a sample of exhaled breath.

Back in 2008, Jun Ye and colleagues at JILA demonstrated the world's first frequency comb breathalyzer, which measured the absorption of light in the near-infrared part of the optical spectrum. In 2021 they achieved a thousandfold improvement in detection sensitivity by extending the technique to the mid-infrared spectral region, where molecules absorb light much more strongly. This enables some breath molecules to be identified at the parts-per-trillion level where those with the lowest concentrations tend to be present.

The added benefit to this study was the use of machine learning. Machine learning—a form of artificial intelligence (AI)—processes and analyzes a massive, complex mélange of data from all the breath samples as measured by 14,836 comb "teeth," each representing a different color or frequency to create a predictive model to diagnose disease.

"Molecules increase or decrease in their concentrations when associated with specific health conditions. Machine learning analyzes this information, identifies patterns and develops reliable criteria we can use to predict a diagnosis," said Qizhong Liang, a graduate student in the Jun Ye group, who is lead author of a new paper presenting the findings.

JILA is jointly operated by the National Institute of Standards and Technology (NIST) and the University of Colorado Boulder (CU Boulder). The research was conducted on breath samples collected from 170 CU Boulder students and staff from May 2021 to January 2022. Approximately half of the volunteers tested positive for COVID-19 with standard PCR tests. The other half of the subjects tested negative. The young study group had a median age of 23 years old, and all were above 18 years old. The general campus population was more than 90% vaccinated.

"I do think that this comb technique is superior to anything out there," NIST/JILA Fellow Jun Ye said. "The basic point is not just the detection sensitivity, but the fact that we can generate a far greater amount of data, or breath markers, really establishing a whole new field of 'comb breathomics' with the help of AI. With a database, we can then use it to search and study many other physiological conditions for human beings and to help advance the future of healthcare."

The JILA comb breathalyzer method demonstrated excellent accuracy for detecting COVID by using machine learning algorithms on absorption patterns to predict SARS-CoV-2 infection. H_2O (water), HDO (semi-heavy water), H_2CO (formaldehyde), NH_3 (ammonia), CH3OH (methanol), and NO_2 (nitrogen dioxide) were identified as discriminating molecules for detection of SARS-CoV-2 infection.

The team measured the accuracy of their results by creating a data graph comparing their predictions of COVID-19 against the PCR test results (which, it should be noted, have high but not perfect accuracy). On the graph, they computed a quantity known as the "area under the curve" (AUC). An AUC of 1, for example, would be expected for perfectly discriminating between ambient air and exhaled breath. An AUC of 0.5 would be expected for making random guesses on whether the individuals were born on odd or even months. The researchers measured an AUC of 0.849 for their COVID-19 predictions. An AUC of 0.8 or greater for medical diagnostic data is considered "excellent" accuracy.

In the future, the researchers could further increase the accuracy by expanding the spectral coverage, analyzing the patterns with more powerful AI techniques, and measuring and analyzing additional molecules, which could include the SARS-CoV-2 virus itself. Researchers would need to build a database of the specific IR colors absorbed by the virus (its spectral "fingerprint") to potentially measure viral concentrations in the breath.

The researchers also identified significant differences in breath samples based on tobacco use and a variety of gastrointestinal symptoms such as lactose intolerance. This suggests broader capability of the technique for diagnosing different sets of diseases.

The researchers plan further studies to try to diagnose other conditions such as chronic obstructive pulmonary disease, the third-leading cause of death worldwide according to the World Health Organization. The researchers have also recently boosted the comb breathalyzer's diagnostic power by expanding the spectral coverage to detect additional molecules. They plan to employ additional AI approaches such as deep learning to improve its disease-detection abilities. Efforts are already under way to miniaturize and simplify the technology to make it portable and easy to use in hospitals and other care settings.

Ye said there is interest from the medical community in seeing the comb breathalyzer developed further and commercialized. Approval by the U.S. Food and Drug Administration (FDA) would be needed before the technology could be used in medical settings.

The most prevalent analytical technique in breath research now is gas chromatography combined with mass spectrometry, which can detect hundreds of exhaled molecules but works slowly, typically requiring tens of minutes. Its use of chemical process also unavoidably alters breath components and presents analytical challenges to identify breath profiles accurately. Frequency comb technology measures breath molecules in a non-destructive and real time manner and can promote a more accurate and repeatable determination of exhaled breath contents.

More information: Qizhong Liang et al, Breath analysis by ultra-sensitive broadband laser spectroscopy detects SARS-CoV-2 infection, *Journal of Breath Research* (2023). DOI: 10.1088/1752-7163/acc6e4