



Novel, accurate pathogen sensors for fast detection of SARS-CoV-2 in the aerosol in seconds for a breathalyzer platform

Xiaoling Shi^a, Pardis Sadeghi^b, Nader Lobandi^b, Shadi Emam^b, Seyed Mahdi Seyed Abrishami^b, Isabel Martos-Repath^b, Natesan Mani^b, Mehdi Nasrollahpour^b, William Sun^b, Stav Rones^b, Joshua Kwok^b, Harsh Shah^b, Joseph Charles^b, Zulqarnain Khan^b, Sheree Pagsuyoin^c, Akarapan Rojjanapinun^c, Ping Liu^d, Jeongmin Chae^e, Maxime Ferreira Da Costa^e, Jianxiu Li^e, Xin Sun^f, Mengdi Yang^f, Jiahe Li^f, Jennifer Dy^b, Jennifer Wang^d, Jeremy Luban^d, ChingWen Chang^d, Robert Finberg^d, Urbashi Mitra^e, Sydney Cash^g, Gregory Robbins^g, Cole Hodys^a, Hui Lu^a, Patrick Wiegand^h, Robert Rieger^h, Nian X. Sun^{a,b,*}

^a Winchester Technologies, LLC, Burlington, MA, USA

^b Electrical & Computer Engineering, Bioengineering & Chemical Engineering, W.M. Keck Laboratory for Integrated Ferroics, Northeastern University, Boston, MA, USA

^c Civil and Environmental Engineering, Center for Pathogen Research & Training (CPRT), U. of Massachusetts, Lowell, MA, USA

^d University of Massachusetts Medical School, Worcester, MA, USA

^e Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, USA

^f Bioengineering, Northeastern University, Boston, MA, USA

^g Massachusetts General Hospital, Boston, MA, USA

^h Networked Electronic Systems, University of Kiel, Kiel, Germany

ARTICLE INFO

Keywords:

Pathogen sensors
Biomarker
SARS-CoV-2
Spike protein
Breathalyzer
VOC

ABSTRACT

Rapid and accurate detection of the pathogens, such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) for COVID-19, is critical for mitigating the COVID-19 pandemic. Current state-of-the-art pathogen tests for COVID-19 diagnosis are done in a liquid medium and take 10–30 min for rapid antigen tests and hours to days for polymerase chain reaction (PCR) tests. Herein we report novel accurate pathogen sensors, a new test method, and machine-learning algorithms for a breathalyzer platform for fast detection of SARS-CoV-2 virion particles in the aerosol in 30 s. The pathogen sensors are based on a functionalized molecularly-imprinted polymer, with the template molecules being the receptor binding domain spike proteins for different variants of SARS-CoV-2. Sensors are tested in the air and exposed for 10 s to the aerosols of various types of pathogens, including wild-type, D614G, alpha, delta, and omicron variant SARS-CoV-2, BSA (Bovine serum albumin), Middle East respiratory syndrome-related coronavirus (MERS-CoV), influenza, and wastewater samples from local sewage. Our low-cost, fast-responsive pathogen sensors yield accuracy above 99% with a limit-of-detection (LOD) better than 1 copy/μL for detecting the SARS-CoV-2 virus from the aerosol. The machine-learning algorithm supporting these sensors can accurately detect the pathogens, thereby enabling a new and unique breathalyzer platform for rapid COVID-19 tests with unprecedented speeds.

1. Introduction

The COVID-19 pandemic has claimed the lives of roughly 7 million people. Rapid and timely COVID-19 tests can effectively mitigate the pandemic, increase the chance of survival, and limit its side effects. Since the outbreak of the COVID-19 pandemic in 2019, different

methods have been developed and commercialized for COVID-19 tests. RT-PCR (reverse transcription polymerase chain reaction) tests have been the gold standard for a medical diagnosis of COVID-19, but they are slow with hours to days turnaround time and require expensive instrumentation and skilled staff to carry out the tests (Vandenberg et al., 2021; Wang et al., 2021; Ayankojo et al., 2022; Farsaeivahid et al.,

* Corresponding author. Winchester Technologies, LLC, Burlington, MA, USA.

E-mail address: n.sun@northeastern.edu (N.X. Sun).

<https://doi.org/10.1016/j.biosx.2023.100369>

Received 30 January 2023; Received in revised form 22 May 2023; Accepted 8 June 2023

2590-1370/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2022). Antibody tests in serum based on enzyme-linked immunosorbent assay are cheaper and faster than PCR tests. But they are not appropriate for early-stage diagnosis (Rump et al., 2021; Raziq et al., 2021) as antibodies cannot be detected until 10–14 days after symptoms begin to appear (Peto et al., 2021; Rump et al., 2021). Antigen tests based on lateral flow immunoassay take 10–30 min, but they have limited sensitivity (Peto et al., 2021). Therefore, there is an urgent need to develop new pathogen sensors which can detect the pathogens, not VOCs, and provide test results in seconds (not minutes) for rapid detection of viral and bacterial pathogens such as SARS-CoV-2.

Recently, different laboratory-based rapid detection methods for SARS-CoV-2 viruses have emerged. Porte et al. (2020) developed a novel antigen-based rapid test for SARS-CoV-2 in respiratory samples based on fluorescent immunochromatography. Zamzami et al. (2022) designed a carbon-nanotube field-effect transistor (CN-FET)-based biosensor for fast detection (2–3 min) of SARS-CoV-2 surface spike protein S1, with a sensitivity of 4.12 fg/mL. Mautner et al. (2020) proposed a reverse transcription loop-mediated isothermal amplification (RT-LAMP) on heat-inactivated samples to directly detect SARS-CoV-2, which is 12 times faster and 10 times cheaper than RT-PCR COVID-19 tests. Ventura et al. (2020) developed a colorimetric sensor using gold nanoparticles for SARS-CoV-2 surface spike protein detection, which led to a threshold cycle value of $C_t = 36.5$ as the limit of detection of the biosensor in terms of RT-PCR cycle threshold. Huang et al. (2020) designed a double-antibody sandwich plasmonic resonance immunoassay for SARS-CoV-2 pseudo virus detection using an Au nano-cup array chip and gold nanoparticles. Aithal et al. (2022) reported an aptamer-functionalized gold nanoparticle-based sensor that can detect 16 nM spike protein and 3.54×10^3 genome copies/ μL of inactivated SARS-CoV-2. The main problem with these detection methods is that they detect the SARS-CoV-2 pathogens dispersed in a liquid. As such, they require liquid sample preparation, which is time-consuming. A breathalyzer platform that can detect the SARS-CoV-2 pathogen in the aerosol taken from the exhaled breath is more advantageous for COVID-19 screening or diagnosis, particularly if it can deliver test results in less than a minute.

The major challenge for such a breathalyzer for COVID-19 tests is the lack of pathogen sensors to detect SARS-CoV-2 from the aerosol. Conventional breathalyzers can detect small biomarker VOCs with high equilibrium partial pressure at room temperature (Das et al., 2016). Pathogens such as viruses, bacteria, and fungi have much larger molecular weights and are transmitted in liquid droplets or aerosol particles (Gralton et al., 2011). Most test methods for pathogens in aerosols (e.g., gas chromatography-mass spectroscopy, GC-MS) are time-consuming because they rely on the aerosol collection and conversion into a liquid form for conventional liquid assay measurement techniques such as PCR (Aithal et al., 2022; Li et al., 2021). The only FDA (U.S. Food and Drug Administration)-approved COVID-19 breathalyzer has a desktop-sized GC-MS tester for *in vitro* qualitative identification of five VOCs from the ketone and aldehyde families linked with SARS-CoV-2 infection in patients' exhaled breath in 3 min utilizing a 115 VAC power source (U.S. Food and Drug Administration, 2022; Rubin, 2022). Since a unique set of biomarker VOCs for COVID-19 has not been reported, which may also change in response to the different symptoms associated with different variants of SARS-CoV-2 infection, it is challenging for this VOC-based breathalyzer to make an accurate diagnosis, particularly for new SARS-CoV-2 variants. A better idea is to directly detect pathogens such as SARS-CoV-2 (instead of small-molecule VOCs) from the aerosols in exhaled breath, as it leads to much more accurate test results.

Since 2017, we have been working on functionalized molecularly imprinted polymers (MIP)-based sensors for biomarker VOCs for the diagnosis of different diseases, such as Alzheimer's disease (Emam et al., 2018), lung cancer (Emam et al., 2022). As soon as COVID was declared a pandemic in early 2020, we have started to work on the functionalized molecularly-imprinted polymers (MIP)-based electrochemical sensors for detecting the SARS-CoV-2 pathogens from the aerosol for a COVID-19 breathalyzer. Our initial efforts used the full-length spike proteins of the SARS-CoV-2 as the template molecules for making these MIP sensors, which led to the high sensitivity of the sensors but relatively poor specificity. Since full-length spike proteins have molecular weights in the range of 180–200k Da (Huang et al., 2020), which are too

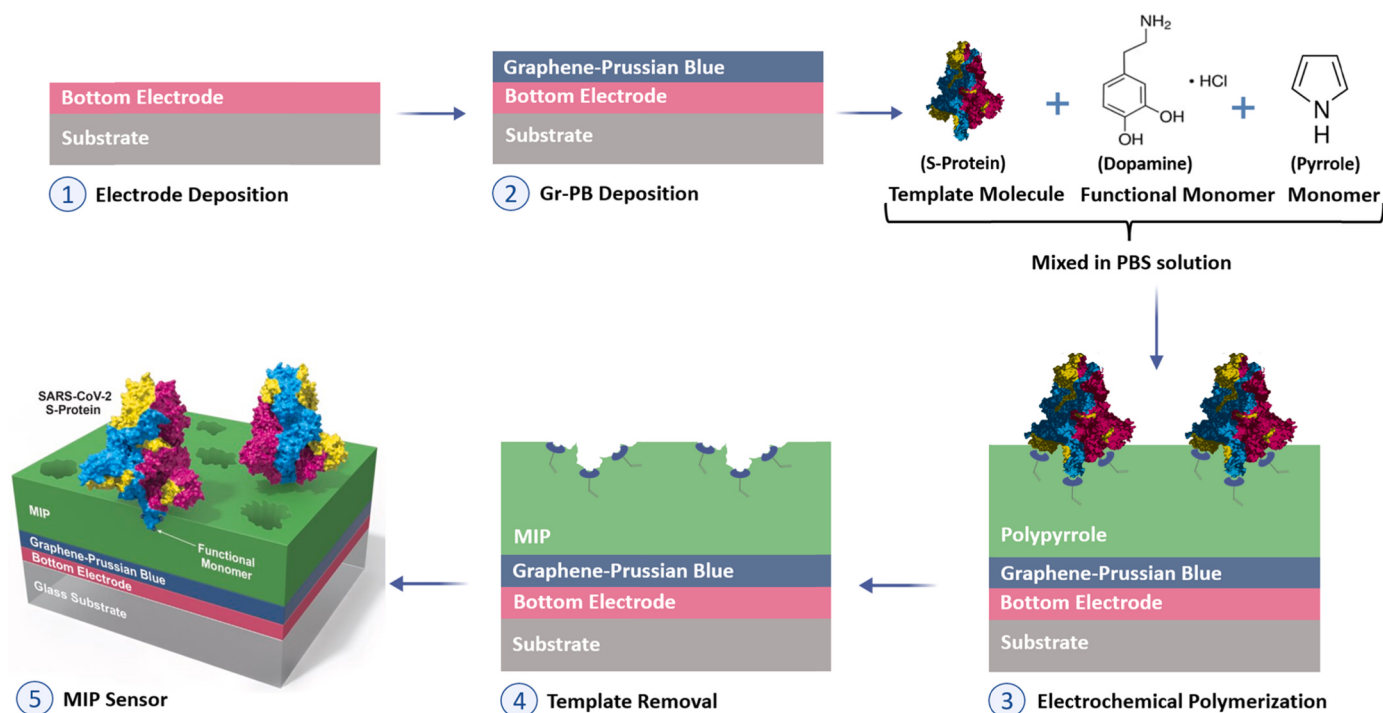


Fig. 1. Schematic of the fabrication process flow and test of the pathogen sensor with functional monomers and template molecules (SARS-CoV-2 s-proteins) forming an artificial antibody (lock) to detect SARS-CoV-2 pathogens (key) with high specificity.

large for MIP sensors, we tried to use the S1 subunit of the spike proteins of the SARS-CoV-2 with molecular weights of nearly 78 kDa (Zamzami et al., 2022). At the same time, we used functionalized MIP sensors by using 1-pyrenebutyric acid N-hydroxy succinimide ester (PBSE) and cysteamine to bind the S1 proteins before using the functionalized S1 proteins as template molecules. This effort led to pathogen sensors for detecting SARS-CoV-2 from the aerosol with significantly improved sensitivity and specificity (Sun et al., 2021). To further improve the sensitivity and specificity, we used the RBD subunit of the spike proteins of the SARS-CoV-2 which further reduced the molecular weight to nearly 26 kDa and successfully achieved ultra-high sensitivity by new functional monomer of dopamine. With these background efforts in mind, we report, for the first time, on developing highly accurate pathogen sensors, new test methods and algorithms for ultrafast detection of SARS-CoV-2 in aerosols. The SARS-CoV-2 pathogen sensors developed in this work are based on the omicron-variant RBD spike proteins and are highly accurate (>99%) for different variants of SARS-CoV-2. The sensor exhibits a fast response and short recovery time, yielding test results in less than 30 s. The proposed approach is an electrochemical sensor that incorporates molecularly-imprinted polymer with functionalization leading to high accuracy. The sensor works on the premise that the ohmic resistance of the device is altered when COVID-19 pathogens are present in the test aerosol sample. In what follows, we give some details about the structure and working principles of this sensor, together with some experimental data.

2. Materials and methods

2.1. Materials

The readers are referred to Supplementary Material (S1) for Research Methodology and Materials, which incorporates design details and fabrication steps. Here, we present our sensor design and fabrication, and the methodology for carrying out a rapid COVID-19 screening or diagnostic test.

2.2. Design and fabrication

Our novel electrochemical pathogen sensors are based on functionalized molecularly imprinted polymers (MIP) and Graphene/Prussian blue on a substrate surface to create synthetic recognition sites in polymeric matrices that are complementary to the targeted template molecules or recognition sites of pathogens in terms of size, shape, atomic groups, and spatial arrangement. A thin layer of the polymer matrix is electrochemically deposited around the pathogens' template molecules or recognition sites, such as the receptor binding domain (RBD) spike proteins (s-protein) of the SARS-CoV-2. The template molecules are then washed away after polymerization, leaving imprinted cavities of the exact size and shape of the template molecule in the polymer matrix. The imprinted cavities in the sensor act as synthetic antibodies. As a result, the sensor only detects the originally-targeted template molecules or pathogens due to the matched shape and size and the strong hydrogen bonding between the cavities on the sensors and the template molecules, leading to extremely high specificity. The binding of the template molecules and the imprinted cavities in the sensor polymer (MIP) layer leads to reduced mobility of charge carriers in the polymer layer and increased sensor ohmic resistance that can be readily measured by a digital multimeter (Zarejousheghani et al., 2021). The steps used for sensor design, fabrication, and testing approach are shown in Fig. 1.

The first step of the fabrication process is the deposition of a metal electrode on a substrate such as Si or glass, and the second step is the deposition of graphene-Prussian blue on the metal electrodes. After deposition of the Graphene-Prussian blue on the surface of a metal electrode, a polypyrrole layer is formed on the Graphene Prussian blue through cyclic voltammetry-induced electropolymerization in a pyrrole-PBS solution. Functional monomers are carefully selected and added during the electropolymerization process to improve the binding between the atom groups between the template molecules and the specificity of the sensors. The selective recognition abilities of the imprinted polymers are mainly due to the formation of a complex between the analyte target and the functional monomers during the pre-

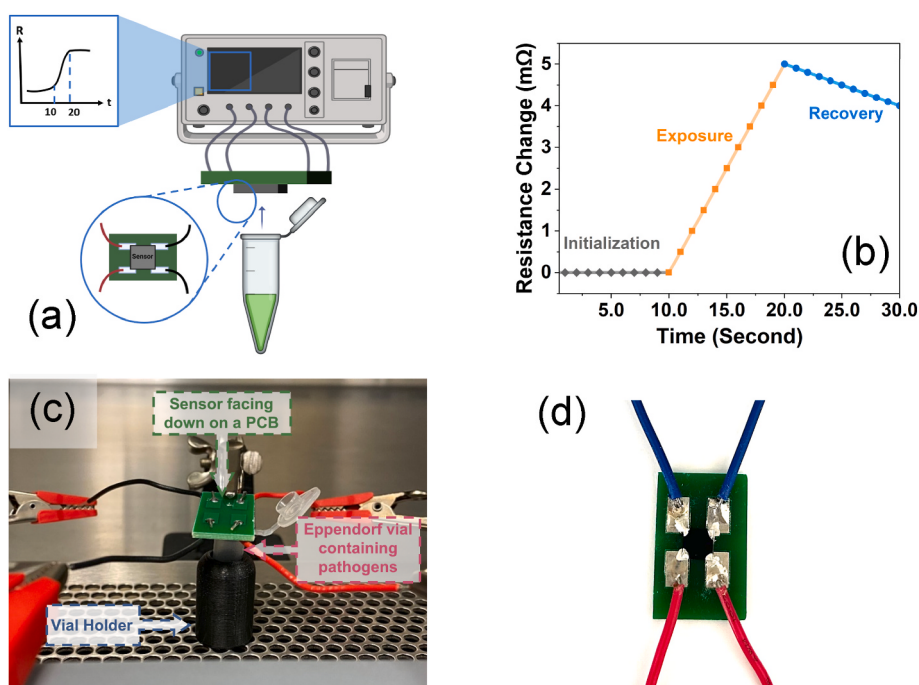


Fig. 2. Schematic depicting the 10s-10s-10s testing method: (a) Sensor resting in the middle of a PCB with four leads for resistance measurement. (b) An idealized test result, (c) Testing setup and its peripherals. (d) Four-lead connections. The sensor is tested with its faced down while an open Eppendorf vial (containing the pathogens to be tested) is placed at 2–3 mm beneath the sensor. During the test, the sensor resistance is measured by an Ohmmeter.

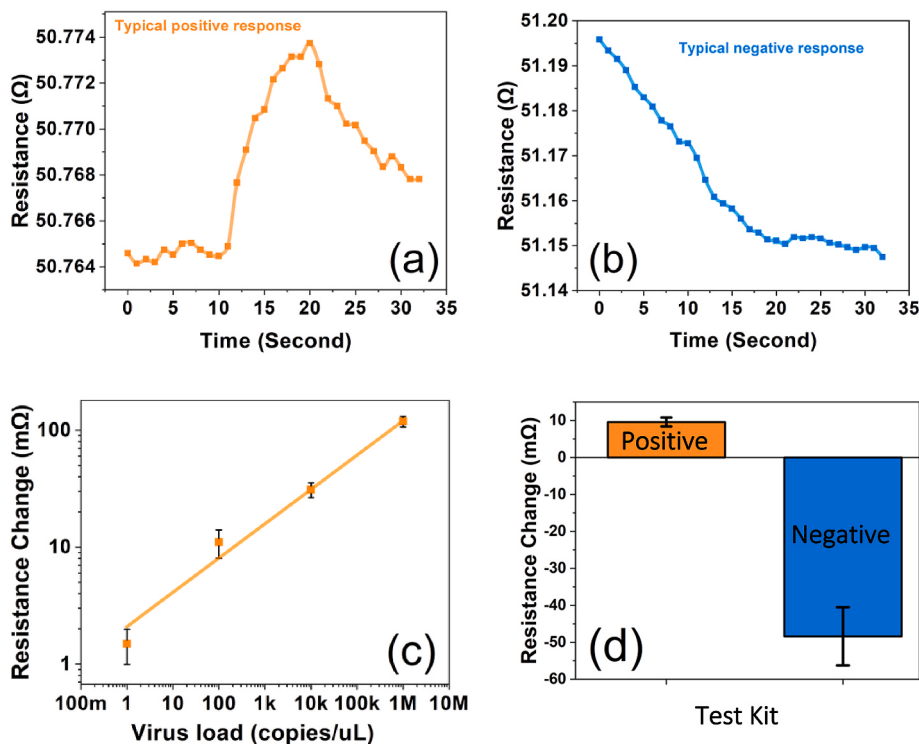


Fig. 3. (a) A typical positive test case with sensor exposed to the aerosol of Delta variant Spike Protein (10 pg/ μ L) during the 10s-10s-10s test method. A similar trend was observed for all fabricated sensors. Note that a resistance change of roughly 1m Ω is the resistance limit of detection which corresponds to a limit of detection (LOD) better than 1 copy/ μ L for SARS-CoV-2 virus; (b) A typical negative test case with sensor exposed to the aerosol of bovine serum albumin protein (BSA); (c) Effect of the virus load on resistance change for a positive test case with the sensor exposed to Delta variant Spike Protein; (d) Repeatability report for a positive test case of delta variant spike protein (10 pg/ μ L) with the mean resistance-change of +9.59 m Ω . This plot also shows repeatability of a negative test case for bovine serum albumin protein (BSA) control with the mean resistance-change of -48.4 m Ω . The relative standard deviation (RSD) shows that, compared with the mean, variance of our data is 12.6% for the positive test case and 16.2% for the negative test case.

polymerization step (Vasapollo et al., 2011). The choice of a suitable functional monomer is based on its ability to establish a good interaction with the functional groups of the template molecule in a covalent or non-covalent way. The choice of functional monomer is critical to achieving a high-accuracy electrochemical pathogen sensor. We chose dopamine as the functional monomer for our sensors based on its prior success in the specific detection of bovine hemoglobin (Li et al., 2015) as well as dopamine's ability to self-polymerize into thin adherent poly-dopamine (PDA) films, which can then be coated on various organic and inorganic substrates (Li et al., 2015). As a functional monomer, dopamine has led to pathogen sensors with the highest accuracy >99% for different variants of SARS-CoV-2, compared to other functional monomers we have used, such as methacrylic acid (MAA). We have, therefore, focused our efforts on dopamine-based sensors in this work (Palladino et al., 2019). After the electropolymerization process, the sensor was washed in ethanol to remove the template molecules and left to dry in the air before electrodes were added for sensor testing. All process steps were carefully optimized. Details of the sensor fabrication process are described in the **Supplemental Materials: S1**.

2.3. New test strategy

One of the main challenges of directly detecting pathogens in the aerosol (i.e., with no need to convert the aerosol into liquid) is that there is no known methodology for testing the pathogens in the aerosol. Commercially available breath analyzers (breathalyzers) used for blood alcohol content (BAC) tests can provide the test result in a couple of seconds by measuring the ethanol from the exhaled breath (Jones, 2016). Recently, we demonstrated that electrochemical sensors based on MIP are effective for measuring the VOCs from exhaled breath with response time on the order of 5–10 min (Adams et al., 2019; Emam et al., 2018; Sun et al., 2022; Emam et al., 2020). As mentioned earlier, the FDA-approved breathalyzer for COVID-19 tests needs roughly 3 min for a test as compared with the antigen tests, which take 5–30 min (Abusrewil et al., 2021). An ideal test method for a COVID-19 breathalyzer should render accurate test results in seconds not minutes. Herein we report on a new method (labelled 10s-10s-10s method) to test the

viability of our novel pathogen sensors. Our testing approach is based on the pathogen binding to the sensor, which leads to immobilization of the carriers in the semiconducting molecularly imprinted polymer layer thereby reducing carrier mobility while significantly increasing its ohmic resistance. To test the sensors' performance, based on our new 10s-10s-10s method, we measure the sensor resistance for 10 s of sensor stabilization, then 10 s of sensor exposure (i.e., with the sensor exposed to the specific aerosol), and finally another 10 s for sensor recovery. The first 10 s allows the sensor to stabilize and sets the baseline for the noise level. During the next 10 s, the sensor is exposed to aerosol of the test liquid with different virus loads or protein concentrations. The last 10 s of sensor recovery allow us to monitor the sensor resistance change after the sensor exposure. Compared to other test methods, this 10s-10s-10s test method enables rapid COVID-19 test result in just 30 s without sacrificing the accuracy. The technique is easy to implement and needs no training.

2.4. Experimental setup

Fig. 2 shows a schematic of the sensor test setup. In the setup, the sensor (which rests on the middle of a PCB board) is faced down with an open Eppendorf vial containing the pathogens placed at 2–3 mm beneath the sensor (see Fig. 2a). An ohmmeter monitors the sensor's ohmic resistance during the operation. Fig. 2b shows an idealized sensor resistance vs. time profile, $R(t)$, obtained this way. For a positive test, the resistance comprises: (1) an initialization 10-s phase, where the resistance is almost constant; (2) an exposure 10-s phase, where the resistance linearly increases with time until it reaches a maximum; and (3) a recovery 10s phase, where the signal drops with time because there is no longer any exposure. The device relies on a 4-point measurement technique (see Fig. 2c and d), which yields more accurate results than the 2-point measurements.

Table 1

A comparison between the human-brain decision with the outputs of three different theoretical classifying methods in terms of sensitivity, specificity, and accuracy (obtained based on the sensor's R(t) profiles) for a sensor exposed to a positive test kit.

Omicron-Variant Pathogen Sensors			
Method	Sensitivity	Specificity	Accuracy
Human Brain Decision	98.41%	100%	99.26%
Wavelet Method	92.10%	90.30%	91.10%
Deep-Learning Method	95.20%	90.30%	92.60%
Curve-Fitting Method	95.23%	100%	97.78%

3. Results and discussions

3.1. Experimental results

Fig. 3 shows a typical positive test result obtained using the 10:10:10 method for the Delta variant Spike Protein (10 pg/mL). As is seen in Fig. 3a, due to the random noise (which is nearly 1 m Ω RMS) the resistance is not exactly equal to zero during the initialization phase. During the exposure phase, the resistance sharply increases with time until it reaches its maximum at $t = 20$ s, at which time the exposure is cut off. The resistance then drops with time during the recovery phase. The response becomes totally different when the sample does not contain the pathogen. Fig. 3b shows the negative test result for bovine serum albumin protein (BSA), which contains no pathogen. As is seen in Fig. 3b, for a negative test the resistance profile is a decaying function of time, on a time-averaged basis. In fact, negative response is mainly dominated by the noise during all three phases. The limit of detection of these pathogen sensors can be estimated when the resistance change is equal to the noise level at 1 m Ω leading to a signal-to-noise ratio of 1 or, equivalently, 0 dB. The limit of detection for the sensor shown in Fig. 3 is between 0.1 and 1 copies/ μ L (i.e., 0.1–1 fg/mL), which is comparable to the conventional RT-PCR (Cheong et al., 2020). It is important to note that, for the positive test cases, the load plays a key role on the ohmic resistance. It is speculated that the resistance-change (ΔR) increases with the pathogen loads of the gamma-ray inactivated wild-type SARS-CoV-2 virus. The log-log plot in Fig. 3c shows that this is indeed the case. In fact, based on the data presented in this figure, ΔR increases with the load in a nonlinear fashion. The best fit turned out to be of the power-law form: $\Delta R = k \times (\text{virus_load})^n$ where the coefficient (k) and the exponent (n) are found to be equal to 0.2748 ± 0.1064 and 0.2935 ± 0.0209 , respectively—the uncertainties associated with (k,n) have been obtained from linear regression analysis on the log-transformed data. The error bars in Fig. 3c are a clear indication of the errors associated with the electrical measurements. We also tested our sensors with RT-qPCR wastewater samples, the readers are referred to Supplementary Material (S2) for the results.

It needs to be mentioned that the pathogen sensors (made with the RBD spike-proteins of omicron-variant SARS-CoV-2 for detecting different variants of SARS-CoV-2 pathogens) include gamma-ray inactivated wild-type SARS-CoV-2 virion particles and delta-variant spike proteins, and they all showed a sensitivity of 98.41%, specificity of 100% and accuracy of 99.26%. Detailed test results for each of the sensors are listed in Supplementary Material: S2. All pathogen sensors were carried out on a custom-designed test kit with 15 samples, including 6 positive-labelled and 9 negative-labelled controls. The current test kit shows 62 TP (True Positive), 1 FN (False Negative), 72 TN (True Negative), and 0 FP (False Positive) for the omicron-variant pathogen sensors. Statistical parameters of our sensor has been shown in Table 1. The positive test kits include Gamma-ray inactivated wild-type SARS-CoV-2 virion particles with different virus loads of 1, 10^2 , 10^4 , and 10^6 copies/ μ L, and the spike proteins of delta-variant SARS-CoV-2 with a concentration of 1 and 10 pg/ μ L. Negative controls for the test kit will include Vero E6 cell lysate, BSA, PBS, influenza A/B, MERS, and SARS (or SARS-1) with typical concentrations of 0.1–10 pg/ μ L. We have tested hundreds of such sensors. These pathogen sensors were also tested at different temperatures from 0 $^{\circ}$ C up to 50 $^{\circ}$ C in an environmental containing contaminant gases (water vapor, CO₂, ammonia, etc.) to simulate real-world breath tests. Our sensors were found to be insensitive to temperatures and environmental contaminants. Based on the obtained experimental data, sensitivity, specificity, and accuracy of the sensors made with omicron-variant RBD are respectively equal to 98.41%, 100%, and 99.26%.

The SEM investigation of MIP samples was also studied to better understand the sensor surface topography (Chang and Kuo, 1993; Yu et al., 2011). We deposited a thin layer of roughly 3 nm of Gold Palladium on the sensor's surface to take a better SEM image. Fig. 4 shows SEM images of the sensor before/after electrodeposition, and also when the template is removed. As is seen in this figure, the surface topographies are significantly different from each other. The SEM images of the Gr-PB layer shows that, after electrodeposition, a polymeric layer has been formed on the top of the Gr-PB layer. On the other hand, Fig. 4c shows that the process of template removal results in the formation of some voids on the surface of the sensor.

The pathogen detection sensors developed in this work and the data presented in Fig. 3 are based on human evaluation of the sensor resistance data. We have also analyzed the capabilities of artificial intelligence (Aster et al., 2018) to detect COVID-19 from the sensor data with no need for visual inspection of the ohmmeter data. Such an approach would enable automated detection with no need for gross human intervention.

3.2. Artificial intelligence methods for COVID-19 diagnosis

In this section, we design and evaluate several detection schemes based on machine learning (Goodfellow et al., 2016). To this end, we

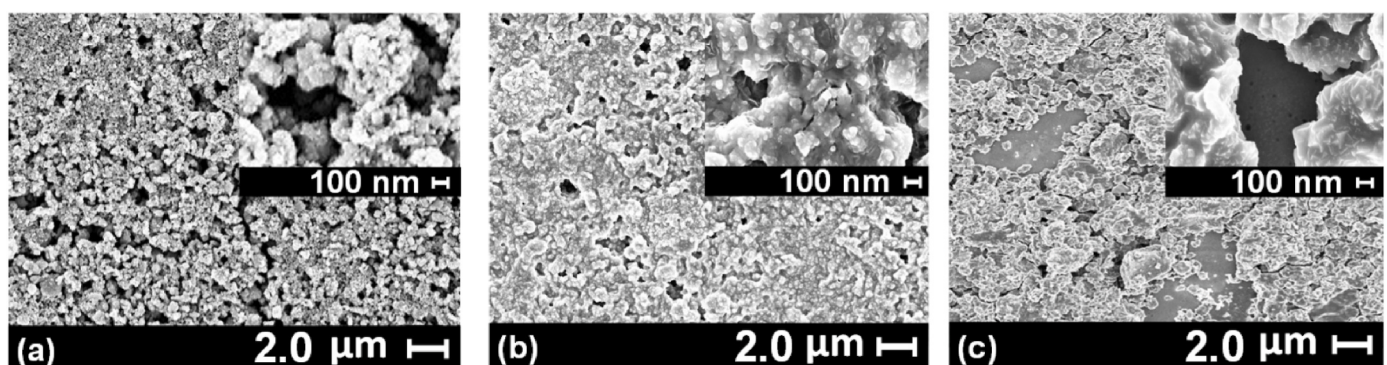


Fig. 4. SEM images of MIP sensor surface: (a) before electrodeposition; (b) after electrodeposition; and (c) after template removal.

rely on an experimental dataset of L time series that were acquired as noted in the prior sections. We denote by $\{x_i\}_{i=1}^L$ the i th experimental time series of the dataset. As seen below, each test set is of length 30, corresponding to a 30-s observation sampled at 1 Hz. Each experimental time series is associated with a known binary label $y_i \in \{0, 1\}$ that takes value 0 or 1 in the absence or presence of SARS-CoV-2 pathogens in the aerosol being tested, respectively. We design a classifier, represented by a function $f: \mathbb{R} \rightarrow \{0, 1\}$ that associates every input time series x of length n with a binary prediction corresponding to inference on the presence of SARS-CoV-2 pathogens in the tested aerosol. An efficient classifier should achieve low empirical risk, such that the number of classified entries in a test dataset of L time series given by $\sum_{i=1}^L R(\hat{y}_i, y_i) = \sum_{i=1}^L R(f(x_i), y_i)$ is small, where the function R equals 0 if $\hat{y}_i = y_i$ and 1 otherwise. In practice, the risk function R can be weighted to balance between the two types of the probability of error: $P[f(x) = 0 | \text{pathogens}]$ and $P[f(x) = 1 | \text{no pathogen}]$. In the sequel, three categories of classifiers are considered. We first consider two model-driven classifiers based on the wavelet decomposition (Dempster et al., 1977; Daubechies, 1990; Rhif et al., 2019). These classical methods rely on prior model assumptions on the structure of the acquired time series. Moreover, they come with the advantage of needing a small number of samples to train the classifier. The wavelet decomposition assumes that there is an underlying deterministic function mapping time to the resistance that linear combinations of wavelet functions can describe. The Gaussian mixture model assumes that the resistance can be described by a random variable with a distribution that is a mixture of Gaussian probability density functions. As these methods are classical, they are described in more detail in the Supplementary Material: S3 (Bottou, 2010; Carbone et al., 2020; Elman, 1990; Esfahani and Sun, 2023; Figueiredo and Jain, 2002; Gers et al., 2000; Graves et al., 2013; Hochreiter and Schmidhuber, 1997; Ji et al., 2021; Lambrou et al., 1998; Mallat, 1999; McLachlan, 1999; Murugappan et al., 2010).

The models are tested on a dataset of 135 experimental time series, among which 63 are positive and 72 are negative. A cross-validation method is adopted to train and test the models (Browne, 2000). This consists of firstly partitioning the time series into five distinct subsets of 27 time series at random. Next, one subset is discarded and left for validation for each of the five training batches, while the four others are retained for training purposes. The accuracy and specificity of the classifier are defined as the mean statistics over the batches. The wavelet-based and Gaussian mixture model classifiers are implemented under MATLAB. The LSTM neural network is implemented on one computer with Nvidia GeForce RTX 3060 Graphical Processing Units and 16 GB of memory. PYTORCH 1.10.0 is employed as the deep learning framework. Consistent with the experiments for the other methods, we use the same dataset for training and testing. Note that, to make a fair comparison with the other methods, there are no samples for validation, i.e., the samples are either for training or testing. For the training process of the deep learning-based method, the learning rate and the momentum are set to 0.01 and 0.9, respectively. The batch size is selected as 32, considering the size of the dataset, and the dropout rate is set to 0.2 for the first three FCLs. The designed network is trained for 2000 epochs. The performance of the two mathematical models (wavelet-based vs. deep-learning-based) in comparison with the decision made by human brain via visual inspection is listed Table 1. For comparison purposes, we have also included curve-fitting results in this table. As is seen in this list, the deep learning-based classification shows a better trade-off between true positive and true negative rates but at a price of more complexity than the supervised learning methods. Still, it should be conceded that in circumstances where the computational training power is limited, the wavelet-based classifier remains a good compromise.

Another look at Table 1 reveals that curve-fitting is competitive with the other methods. In fact, using the same data set of 135 experimental time series, curve-fitting is seen to provide an excellent accuracy of

97.78%. It needs to be mentioned that the curve-fitting approach relies on artificial intelligence's (AI's) capacity to identify COVID-19 pathogen from sensor data. The diagnosis is produced by the algorithm we have developed in Java language using the raw data from the biomarker breathalyzer sensors. We can determine the regression lines for each period using standard deviation and Pearson correlation coefficient. The initialization, exposure, and recovery phases are the three-time windows that make up the sensor response. The algorithm aims to determine if the sensor resistance during the exposure phase was significantly affected to warrant a positive diagnosis as biomarkers increase the sensor's resistance. The algorithm utilizes a total of four conditions to assess whether the aerosol (to the sensor exposed) influenced the sensor resistance that is significant enough or not. The first condition determines if the exposure phase's regression slope is higher than the initialization phase's regression slope. The second condition determines if the exposure phase's regression slope is greater than the recovery phase's regression slope. The third condition determines if the exposure phase's regression slope is greater than zero. The final condition determines whether the exposure phase's range is larger than the initialization phase's range. Conditions one and two check whether the rise in resistance during the exposure phase was greater than that in the initialization and recovery phases, indicating an impact created by the substrate on the sensor. The third condition ensures that the magnitude of change is directed in the positive direction during the exposure phase. Finally, the fourth condition determines whether the changes experienced during the exposure phase are significant compared to the changes seen during the initialization phase. If the changes are not more significant, they can be marked down as noise and discarded.

All in all, the analysis described in this sub-section demonstrates that the pathogen sensors are statistically informative for detecting SARS-CoV-2 pathogens in aerosols.

3.3. Viability of our pathogen BioSensor

As earlier mentioned, there are different methods in the market for detecting COVID-19 virus in human body. *Molecular-based methods* use the reverse transcription-polymerase chain reaction RT-PCR to amplify and detect viral RNA from respiratory samples. They have high specificity (>99%) and sensitivity (>95%) for COVID-19 diagnosis, but their limit of detection (LOD) is rather low (Wölfl-Duchek et al., 2022). Figueroa et al. (2021) developed a two-step endpoint RT-PCR assay for SARS-CoV-2 detection with a limit of detection of 20 viral RNA copies/ μL with specificity of 95.8% and a sensitivity of 95.1%. Their method, however, requires specialized equipment and training, and can take several hours to produce test results (Udugama et al., 2020). For reasons like these, other methods have been developed or are under development to improve the speed, accuracy and accessibility of COVID-19 testing. *Antigen-based methods* use immunological assays methods that can detect viral antigens or host antibodies in blood or saliva samples using various platforms such as enzyme-linked immunosorbent assay (ELISA), lateral flow immunoassay (LFIA) or electrochemical biosensors. These methods can provide rapid and point-of-care testing for COVID-19, especially in the early stages of infection (Broughton et al., 2020), but they have lower specificity (85–99%) and sensitivity (23–71%) than molecular methods, and they are prone to false negative results. Their LOD also varies depending on the assay and the sample type. For example, the Abbott BinaxNOW COVID-19 Ag Card has a LOD of 97.1 TCID₅₀/mL for nasal swabs (Mak et al., 2020). *Saliva-based method* uses RT-LAMP to amplify and detect viral RNA from saliva samples. However, they are less sensitive than swab-based tests and can detect SARS-CoV-2 in 45 min. Its LOD is also reported to be 6 copies/ μL (Leleu et al., 2020). In contrast, *CRISPR-based method* uses a CRISPR-Cas12 enzyme that cleaves a reporter molecule when it binds to a specific viral sequence, thereby generating a visible signal on a paper strip. It has high specificity (100%) and sensitivity (95%) for COVID-19 diagnosis, and it can detect SARS-CoV-2 from respiratory swab RNA

Table 2

A comparison between the novel sensor developed in this work with those already available on the market for COVID-19 test.

	Approach	Sensitivity	Specificity	Time	Body Fluids	Vendors
RT-PCR	RNA	95–100%	95–100%	3 h	Respiratory specimens	LabCrop, Roche, etc.
Antibody	IgM, IgG	80–100%	90–100%	15–30 min	Blood	Cellex, etc.
Antigen	RNA	20–71%	85–100%	15–30 min	Blood	UCSD
CRISP-based	RNA	95%	100%	40 min	Respiratory specimens	Sherlock Biosciences
LAMP-based	RNA	97%	100%	30 min	Respiratory specimens	MicrosensDx
Saliva-based	RNA	91%	98%	45 min	Saliva	Yale School of Public Health
Breath-based	GC-MS VOC	91%	99%	5–10 min	Air	Inspect-IR, Breathomix
Our Sensor	RBD S-Protein	98.40%	100%	10 s	Aerosol/Air	Winchester Technologies, LLC

extracts in less than 40 min. Its LOD is reported to be 10 copies/μL. However, this technique requires specialized staff and training (Ganbaatar and Liu, 2021). LAMP-based method uses loop-mediated isothermal amplification (LAMP) to amplify and detect viral RNA from respiratory samples. It has high specificity (100%) and sensitivity (97%) for COVID-19 diagnosis, and it can detect SARS-CoV-2 in 30 min. Its LOD is reported to be 0.2 copies/μL, but it can produce false positives for contaminated samples (Amaral et al., 2021). The breath-based method for COVID-19 detection is a new and developing technology. One such device is the SpiroNose, made by the Dutch company Breathomix. It analyzes the chemical compounds in a person's breath to detect signatures of a coronavirus infection (Arnold, 2022). Another breath-based test that has been authorized by the FDA is the InspectIR COVID-19 Breathalyzer. It uses gas chromatography/mass-spectrometry technique (GC-MS) to separate and identify chemical mixtures and rapidly detect five volatile organic compounds (VOCs) associated with SARS-CoV-2 infection in exhaled breath.

A comparison between our novel breath-based pathogen sensors with other methods discussed above has been shown in Table 2. This table suggests that our novel pathogen biosensor is indeed a viable and competitive option for COVID-19 rapid tests.

4. Conclusion

We have designed and tested a novel pathogen sensor that can directly detect SARS-CoV-2 from the aerosol. The sensors developed in this work respond to the pathogens in the aerosol within 10 s, exhibit a high accuracy of >99%, and can readily distinguish SARS-CoV-2 from influenza viruses, MERS viruses, etc., even in the presence of environmental contaminants. These sensors and the machine-learning algorithm developed by our team have enabled a new breathalyzer platform for rapid COVID-19 screening and diagnosis. A comparison between our pathogen aerosol-based sensor with other methods currently available on the market (see Table 1) demonstrates that our novel sensor is indeed a viable option for rapid COVID-19 tests. Our affordable pathogen sensor is of low-power and this makes it a good option for POC applications. To this should be added the fact that, the sensor is supported by a user-friendly mobile application that analyses the data and provides detailed test results. Work is currently ongoing in our research group to extend the range of applicability of our novel biosensor for detecting other pathogens (viruses, bacteria, or fungi) with minimal human intervention.

CRedit authorship contribution statement

Xiaoling Shi: Writing – original draft, Conceptualization. **Pardis Sadeghi:** Writing – original draft, Conceptualization. **Nader Lobandi:** Writing – original draft. **Shadi Emam:** Conceptualization. **Seyed Mahdi Seyed Abrishami:** Conceptualization. **Isabel Martos-Repath:** Conceptualization. **Natesan Mani:** Conceptualization. **Ping Liu:** Conceptualization. **Xin Sun:** Conceptualization. **Mengdi Yang:** Conceptualization. **Jiahe Li:** Conceptualization. **Jennifer Wang:** Conceptualization. **Jeremy Luban:** Conceptualization. **ChingWen Chang:** Conceptualization. **Robert Finberg:** Conceptualization.

Urbashi Mitra: Conceptualization, Writing – original draft. **Cole Hodys:** Conceptualization. **Hui Lu:** Conceptualization. **Patrick Wiegand:** Conceptualization. **Robert Rieger:** Conceptualization. **Nian X. Sun:** Writing – original draft, Conceptualization, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work has been/was funded by The Trinity Challenge, Carrier Inc., DARPA, NSF 2200221 and 2031142, NIH RADx/BBIC, CBD/DoD, Northeastern COVID-19 Funds, and Upstage Lung Cancer Foundation. The authors would like to thank the reviewers for their constructive and encouraging comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biosx.2023.100369>.

References

- Abusrewil, Z., Alhudiri, I.M., Kaal, H.H., El Meshri, S.E., Ebrahim, F.O., Dalyoum, T., Efrefer, A.A., Ibrahim, K., Elfighi, M.B., Abusrewil, S., Elzagheid, A., 2021. J. Med. Virol. 93, 6512–6518.
- Adams, J.D., Emam, S., Sun, N., Ma, Y., Wang, Q., Shashidhar, R., Sun, N.X., 2019. IEEE Sensor. J. 19, 6571–6577, 2019.
- Arnold, C., 2022. Nat. Biotechnol. 40, 990–993.
- Amaral, C., Antunes, W., Moe, E., et al., 2021. Sci. Rep. 11, 16430, 2021.
- Ayankoj, A.G., Boroznjak, R., Reut, J., Öpik, A., Syrtiski, V., 2022. Sensor. Actuator. B Chem. 353, 131160.
- Aithal, S., Mishriki, S., Gupta, R., Sahu, R.P., Botos, G., Tanvir, S., Hanson, R.W., Puri, I. K., 2022. Talanta 236, 122841.
- Aster, R.C., Borchers, B., Thurber, C.H., 2018. Parameter Estimation and Inverse Problems. Elsevier.
- Bottou, L., 2010. In: Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics Paris France, August 22–27, 2010 Keynote, Invited and Contributed Papers. Physica-Verlag HD, pp. 177–186.
- Browne, M.W., 2000. J. Math. Psychol. 44, 108–132, 2000.
- Broughton, J.P., Deng, X., Yu, G., et al., 2020. Nat. Biotechnol. 38, 870–874.
- Carbune, V., Gonnet, P., Deselaers, T., Rowley, H.A., Daryin, A., Calvo, M., Gervais, P., 2020. Int. J. Doc. Anal. Recogn. 23 (2), 89–102.
- Chang, T., Kuo, C.C., 1993. IEEE Trans. Image Process. 2 (4), 429–441.
- Cheong, J., Yu, H., Lee, C.Y., Lee, J.U., Choi, H.-J., Lee, J.-H., Lee, H., Cheon, J., 2020. Nat. Biomed. Eng. 4, 1159–1167.
- Daubechies, I., 1990. IEEE Trans. Inf. Theor. 36, 961–1005, 1990.
- Das, S., Pal, S., Mitra, M., 2016. Med. Biol. Eng. Comput. 36, 605–624.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. J. Roy. Stat. Soc. B 39, 1–22.
- Elman, J.L., 1990. Cognit. Sci. 14 (2), 179–211.
- Emam, S., Adedoyin, A., Geng, X., Zaeimbashi, M., Ekenseair, A., Adams, J., Podlaha-Murphy, E., Sun, N.X., 2018. J. Sens. 2018, 3437149.

- Emam, S., Nasrollahpour, M., Colarusso, B., Cat, X., Grant, S., Kulkarni, P., Ekenseair, A., Gharagouzloo, C., Fris, C.F., Sun, N.-X., 2020. *Diagn. Assess. Dis. Monit.* 12, e12088, 2020.
- Emam, S., Nasrollahpour, M., Allen, J.P., He, Y., Hussein, H., Shah, H.S., et al., 2022. *Biomed. Microdevices* 24 (4), 41.
- Esfahani, I.C., Sun, H., 2023. *Sensor Actuator Phys.* 350, 114121.
- Farsaeivahid, N., Grenier, C., Nazarian, S., Wang, M.L., 2022. *Sensors* 23 (1), 433.
- Figueroa, S., Freire-Paspuel, B., Vega-Marino, P., et al., 2021. *Sci. Rep.* 11, 21658.
- Figueiredo, M.A.T., Jain, A.K., 2002. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 381–396.
- Gers, F.A., Schmidhuber, J., Cummins, F., 2000. *Neural Comput.* 12 (10), 2451–2471.
- Graves, A., Mohamed, A.R., Hinton, G., 2013. May). In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp. 6645–6649.
- Ganbaatar, U., Liu, C., 2021. *Front. Cell. Infect. Microbiol.* 11, 663949.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT press.
- Gralton, J., Tovey, E., McLaws, M.-L., Rawlinson, W.D., 2011. *J. Infect.* 62, 1–13.
- Hochreiter, S., Schmidhuber, J., 1997. *Neural Comput.* 9 (8), 1735–1780.
- Huang, Y., Yang, C., Xu, X.-f., Xu, W., Liu, S.-W., 2020. *Acta Pharmacol. Sin.* 41, 1141–1149.
- Jones, A.W., 2016. Alcohol: breath analysis. In: Payne-James, Jason, Byard, Roger W. (Eds.), *Encyclopedia of Forensic and Legal Medicine*, second ed. Elsevier, pp. 119–137.
- Ji, S., Ran, R., Chiniforooshan Esfahani, I., Wan, K.T., Sun, H., 2021. November). In: ASME International Mechanical Engineering Congress and Exposition, vol. 85598. American Society of Mechanical Engineers, V005T05A075.
- Lambrou, T., Kudumakis, P., Speller, R., Sandler, M., Linney, A., 1998. May). In: Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181), vol. 6. IEEE, pp. 3621–3624.
- Leleu, O., Vincent, G., Auquier, M., et al., 2020. *Respir Med Res* 77, 95–99, 2020.
- Li, L., Fan, L., Dai, Y., Kan, X., 2015. *Microchim. Acta* 182, 2477–2483, 2015.
- Li, M., Wang, L., Qi, W., Liu, Y., Lin, J., 2021. *Micromachines* 12, 798, 2021.
- Mautner, L., et al., 2020. *Virology* 17, 1–14, 2020.
- Mallat, S., 1999. *A Wavelet Tour of Signal Processing*. Elsevier.
- Mak, G.C., Lau, S.S., Wong, K.K., et al., 2020. *J. Clin. Virol.* 2020.
- McLachlan, G.J., 1999. *Resonance* 4 (6), 20–26.
- Murugappan, M., Ramachandran, N., Sazali, Y., 2010. *J. Biomed. Sci. Eng.* 3 (4), 390.
- Peto, T., et al., 2021. *eClinicalMed.* 36, 100924.
- Palladino, P., Bettazzi, F., Scarano, S., 2019. *Anal. Bioanal. Chem.* 411, 4327–4338.
- Porte, L., Legarraga, P., Vollrath, V., Aguilera, X., Munita, J.M., Araos, R., Pizarro, G., Vial, P., Iruretagoyena, M., Dittich, S., Weitzel, T., 2020. *Int. J. Infect. Dis.* 99, 328–333.
- Rhif, M., Ben Abbes, A., Farah, I.R., Martínez, B., Sang, Y., 2019. *Appl. Sci.* 9, 1345.
- Rubin, R., 2022. *JAMA* 327 (19), 1860–1860.
- Rump, A., Risti, R., Kristal, M.L., Reut, J., Syritski, V., Lookene, A., Boudinot, S.R., 2021. *Biochem. Biophys. Res. Commun.* 534, 457–460.
- Raziq, A., Kidakova, A., Boroznjak, R., Reut, J., Öpik, A., Syritski, V., 2021. *Biosens. Bioelectron.* 178, 113029, 2021.
- Sun, N.X., Emam, S., Ekenseair, A.K., 2022. *Molecularly-Imprinted Electrochemical Sensors*, 16/383220. US Patent.
- Sun, N., Emam, S., Nasrollahpourmotlaghzanjani, M., 2021. *Worldwide Patent Application WO2022035462A1*. Provisional Application Serial, 63/065,462.
- Udugama, B., Kadhiresan, P., Kozlowski, H.N., et al., 2020. *ACS Nano*, 2020.
- Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A., Kozlakidis, Z., 2021. *Nat. Rev. Microbiol.* 19, 171–183.
- Vasapollo, G., Del Sole, R., Mergola, L., Lazzoi, M.R., Scardino, A., Scorrano, S., Mele, G., 2011. *Int. J. Mol. Sci.* 12, 5908–5945.
- Ventura, B.D., Cennamo, M., Minopoli, A., Campanile, R., Censi, A.B., Terracciano, D., Portell, G., Velotta, R., 2020. *ACS Sens.* 5, 3043–3048.
- Wang, C.C., Prather, K.A., Sznitman, J., Jimenez, J.L., Lakdawala, S.S., Tufekci, Z., Marr, L.C., 2021. *Science* 373, eabd9149.
- Wölfl-Duchek, M., Bergmann, F., Jorda, A., et al., 2022. *Microbiol. Spectr.* 10 (1), e0202921.
- Yu, G., Sapiro, G., Mallat, S., 2011. *IEEE Trans. Image Process.* 21, 2481–2499.
- Zarejousheghani, M., Rahimi, P., Borsdorf, H., Zimmermann, S., Joseph, Y., 2021. *Sensors* 21, 2406.
- Zamzami, M.A., Rabbani, G., Ahmad, A., Basalah, A.A., Al-Sabban, W.H., Ahn, S.N., Choudhry, H., 2022. *Bioelectrochemistry* 143, 107982.