

http://pubs.acs.org/journal/acsodf

# Artificial Intelligent Olfactory System for the Diagnosis of Parkinson's Disease

Wei Fu, Linxin Xu, Qiwen Yu, Jiajia Fang, Guohua Zhao, Yi Li, Chenying Pan, Hao Dong, Di Wang, Haiyan Ren, Yi Guo, Qingjun Liu, Jun Liu,\* and Xing Chen\*



**ABSTRACT:** *Background*: Currently, Parkinson's disease (PD) diagnosis is mainly based on medical history and physical examination, and there is no objective and consistent basis. By the time of diagnosis, the disease would have progressed to the middle and late stages. Pilot studies have shown that a unique smell was present in the skin sebum of PD patients. This increases the possibility of a noninvasive diagnosis of PD using an odor profile. *Methods*: Fast gas chromatography (GC) combined with a surface acoustic wave sensor with embedded machine learning (ML) algorithms was proposed to establish an artificial intelligent olfactory (AIO) system for the diagnosis of Parkinson's through smell. Sebum samples of 43 PD patients and 44 healthy controls (HCs) from Fourth Affiliated Hospital of Zhejiang University School of Medicine, China, were smelled by the AIO system. Univariate and multivariate methods were used to identify the significant volatile organic compound (VOC) features in the chromatograms. ML algorithms, including support vector machine, random forest



(RF), *k* nearest neighbor (KNN), AdaBoost (AB), and Naive Bayes (NB), were used to distinguish PD patients from HC based on the VOC peaks in the chromatograms of sebum samples. *Results*: VOC peaks with average retention times of 5.7, 6.0, and 10.6 s, respectively, corresponding to octanal, hexyl acetate, and perillic aldehyde, were significantly different in PD and HC. The accuracy of the classification based on the significant features was 70.8%. Based on the odor profile, the classification had the highest accuracy and F1 of the five models with 0.855 from NB and 0.846 from AB, respectively, in the process of model establishing. The highest specificity and sensitivity of the five classifiers were 91.6% from NB and 91.7% from RF and KNN, respectively, in the evaluating set. *Conclusions*: The proposed AIO system can be used to diagnose PD through the odor profile of sebum. Using the AIO system is helpful for the screening and diagnosis of PD and is conducive to further tracking and frequent monitoring of the PD treatment process.

# **1. INTRODUCTION**

Parkinson's disease (PD) is the second most common neurodegenerative disorder of the central nervous system in the world.<sup>1</sup> It has a long course of the disease and requires lifelong treatment, which brings great inconvenience to the patients' work and life. People with PD often suffer from some motor and nonmotor symptoms that vary from patient to patient. Motor symptoms include bradykinesia, tremor, rigidity, and postural instability, and the nonmotor symptoms comprise depression, memory loss, anosmia, constipation, and urinary frequency.<sup>1-3</sup> The prevalence was 65.6-12,500/ 100,000, 537-614/100,000, and 51.3-176.9/100,000 in Europe,<sup>4</sup> North America,<sup>5</sup> and Asia,<sup>6</sup> respectively, and will double by the next 30 years along with population aging as predicted by the Global Burden of Disease (GBD).<sup>7</sup> According to the data from GBD, the disease caused 211,296 deaths and 3.2 million patients with disability-adjusted life-years in 6 million patients in 2016.<sup>8</sup> PD does not have an effective way to cure, but early diagnosis of PD and early medical,

psychological, and social interventions can significantly improve the health-related quality of life, relieve symptoms, and prolong the patients' survival time. Therefore, the standardized diagnosis of PD is very important.<sup>9–11</sup>

Currently, the diagnosis of PD is mainly based on clinical manifestations supplemented by test rating scales, including Hoehn–Yahr stage (H-Y),<sup>12,13</sup> unified PD rating scale (UPDRS),<sup>14,15</sup> nonmotor symptom scale,<sup>16,17</sup> PD-cognitive rating scale,<sup>18,19</sup> and so on. There has been a lack of objective and consistent diagnostic criteria for a long time. Dopamine transporter single-photon emission computed tomography is currently a high-value imaging-assisted diagnosis method,<sup>20,21</sup>

Received:September 13, 2021Accepted:January 11, 2022Published:January 26, 2022





	development cohort			validation cohort		
characteristics	PD $(n = 31)$	HC $(n = 32)$	P-value	PD $(n = 12)$	HC $(n = 12)$	P-value
gender ( <i>n</i> , % ratio)						
male	18 (58.0%)	15 (46.9%)	0.184	6 (50.0%)	5 (41.7%)	0.698
female	13 (41.9%)	17 (53.1%)		6 (50.0%)	7 (58.3%)	
age (median, range)	$64.74 \pm 11.31$	64.81 ± 8.89	0.453	$64.00 \pm 6.86$	$57.97 \pm 16.52$	0.251
BMI	$23.28 \pm 2.26$	$22.06 \pm 1.93$	0.032	$23.82 \pm 2.79$	$23.72 \pm 2.40$	0.310
abuse alcohol (n)	2	0		0	0	
disease process (Avg)	5.23	0		5.25	0	
malignant tumor $(N)$	0	0		0	0	

Table 1. Characteristics of Participants in the Experiments

but this item is expensive and unsuitable for routine use. Moreover, the body fluid (blood and cerebrospinal fluid) biomarkers for the diagnosis of PD are still in the research and verification stage.<sup>22</sup> Pilot studies have shown that the volatile organic compounds (VOCs) in the sebum of PD have a different smell from healthy people, which provides a new idea and method for the diagnosis of PD.

Recently, Nazik described that the increased sebum secretion in PD patients was associated with the increase in the production of yeast and enzymes in the body,<sup>23</sup> and the secretion of hormones may lead to seborrheic dermatitis (SD).<sup>24</sup> SD is considered to be one of the premotor symptoms of PD and has an auxiliary value for diagnosing PD.<sup>25</sup> Trivedi used gas chromatography-mass spectrometry (GC-MS) to prove that there were different VOCs in the sebum of PD and healthy people.<sup>26</sup> These VOCs such as perillic aldehyde and eicosane may change with the increase of sebum secretion, and the interaction between sebum and the yeast of the microbiome can make human skin smelly.<sup>26-29</sup> Besides, Tsuda used GC-MS to analyze the smell of sebum and found that the types and concentrations of VOCs such as dodecane, acetone, and ethyl acetate released in the sebum of PD patients were related to UPDRS part 3,<sup>30</sup> representing the severity of motor in PD. These studies indicated that sebum gas could be used for the detection of PD.

At present, the detection and analysis of VOCs in human sebum in scientific research mainly adopt chromatography. One of the most mainstream methods is to use a combination of GC and general detection technology or gas sensors and electronic noses.<sup>31–34</sup> GC–MS is one of the most common methods, but its bulky size, long analytical time, and high cost may still be unsuitable for clinical use. Fast GC systems, which have been used to detect VOC markers for many diseases, had the characteristics of small size, easy to use, portability, and low cost.<sup>35</sup> These characteristics make it possible to perform pointof-care testing of sebum's smell from PD patients.

In this study, GC combined with a surface acoustic wave (SAW) sensor with embedded machine learning (ML) algorithms was proposed to establish an artificial intelligent olfactory (AIO) system for the diagnosis of Parkinson's through smell. The system has the advantages of being fast, small, easy to operate, portability, and low cost. Experimental data from 31 PD patients and 32 healthy controls (HCs) were obtained by the AIO system. Univariate and multivariate analysis of biomarker features was performed, and ML strategies, including support vector machine (SVM), random forest (RF), k nearest neighbor (KNN), AdaBoost (AB), and Naive Bayes (NB), were used to construct diagnostic biomarker-based models and odor profile-based models. Data from 12 PD and 12 HC were used to evaluate the clinical usage

of the models. The results showed that the AIO system could diagnose PD through the smell of sebum, indicating the potential usage of the AIO system in clinical practice.

# 2. RESULTS

**2.1. Clinical Characteristics.** The VOC data of 31 cases of PD and 32 cases of HC were used to build the Parkinson's odor diagnosis model. The data of 12 cases of PD and 12 cases of HC were used to evaluate the models. The characteristics of patients who participated in the experiment are shown in Table 1.

**2.2. Calibration of AIO.** 2.2.1. Calibration of AIO by Using the Mixed Solution. Figure 1 shows the chromato-



**Figure 1.** Chromatographic frequency response curve of a mixed solution in the AIO system: peak 1 is the internal standard at a concentration of 1.12 mM; peaks 2, 3, 4, 5, 6, 7, and 8 are nonane, decane, undecane, dodecane, tridecane, tetradecane, and pentadecane, respectively, and all VOCs were diluted to parts per million of the original concentration.

graphic frequency response curve of a mixed solution of nonane, decane, undecane, dodecane, tridecane, tetradecane, and pentadecane. The mixed solution was used to prepare standard calibration gases (the details of preparation are shown in Section 5.2.2). The AIO can detect the VOCs in the gas phase. In Figure 1, the horizontal axis is the retention time, which characterized different substances, and the vertical axis is the SAW's response frequency, which characterized the quality of each substance.

2.2.2. Calibration of AIO by Using the Selected Biomarker Reagent. Figure 2 shows spectra for different concentrations (0.025, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 25, and 50 mM) of gasphase octanal (Figure 2A), hexyl acetate (Figure 2B), perillic aldehyde (Figure 2C), and dodecane (Figure 2D) generated by VOC solvents. The clinical range of concentrations was from 0.25 to 2.5 mM.<sup>26,27</sup> The results showed good reproducibility of the AIO system in the detection of these four reagents.



**Figure 2.** Calibrations of AIO by different reagent concentrations of the gas phase (0.025, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 25, and 50 mM): (A) responses of AIO to the concentrations of nine different samples of octanal and the detection spectra of different octanal concentrations in the gas phase; (B) responses of AIO to the concentrations of nine different samples of hexyl acetate and the detection spectra of different hexyl acetate concentrations in the gas phase; and (C) responses of AIO to the concentrations of nine different samples of AIO to the concentrations of nine different samples of AIO to the concentrations of nine different samples of AIO to the concentrations of nine different samples of perillic aldehyde and the detection spectra of different perillic aldehyde concentrations in the gas phase. (D) Responses of AIO to the concentrations of nine different samples of dodecane and the detection spectra of different dodecane concentrations in the gas phase. The linear correlation with red dashed lines represents the fitting with 95% confidence interval.

Figure 2A shows a linear detection range of the system ( $R^2 = 0.9876$ , P < 0.0001) from 0.025 to 50 mM covering the reported octanal concentrations in human sebum. Figure 2B

shows hexyl acetate ( $R^2 = 0.9656$ , P < 0.0001). Figure 2C shows perillic aldehyde ( $R^2 = 0.9099$ , P < 0.0001). Figure 2D shows dodecane ( $R^2 = 0.9919$ , P < 0.0001). The calibration of



Figure 3. ROC curves and box plots of three biomarker features for the discovery cohort. The ROC curve comprehensively considers the characteristics of sensitivity and specificity. Box plots show a comparison of means of log scaled peak intensities of these analytes, where black dots are outliers. In the box plots, the green on the left represents HC, and the orange on the right represents PD patients.

the AIO system indicated that the system had good sensitivity and reproducibility for the quantitative detection of these reagent concentrations.

2.3. Selection of Significant Features Which Classify PD. The data detected by the AIO system were preprocessed (to achieve signal adaptive filtering denoising, baseline correction, drift compensation, normalized). We conducted further analysis and statistical tests on the four extracted features' data to detect the difference between the PD and HC groups. To evaluate the performance of these biomarkers, we used the data of the discovery cohort to perform a nonparametric test. It can be found that octanal and hexyl acetate had significant differences in the Kolmogorov-Smirnov Z test (p < 0.05). Perillic aldehyde was lower in PD samples in the box plots and had significant differences in the Mann-Whitney U test (p < 0.05). However, dodecane was not significantly different in the two tests (p > 0.05). The results of the nonparametric test are shown in Table S2. The area under the curve (AUC) and box plots of the three markers of octanal, hexyl acetate, and perillic aldehyde are shown in Figure 3.

2.4. Model Based on Significant Features Which Classify PD. Table 2 shows the classification effect of the

# Table 2. Model Evaluation Parameters and Clinical Application Evaluation Results

		evaluation parameters			
cohort	accuracy	recall (sensitivity)	precision	F1	specificity
development cohort	0.826	0.982	0.836	0.771	
validation cohort	0.708	0.917			0.500

development cohort's (31 PD and 32 HC) three marker features when the classification model was established. We used 80% for the training set and 20% for the validation set. The accuracy of the model was 82.6%. The F1 of 0.771 also showed that the model has good robustness. Then, we used a blind experiment to diagnose 12 PD and 12 HC, evaluated them with clinical indicators, and found that the accuracy was 70.8%. The sensitivity was 91.7%, which means that the mistake diagnosis rate was less than 10%. However, the specificity effect of the model was 50%. As shown in Figure 4, AUC indicated that the model established by the AIO system had good accuracy.

**2.5. Five Different Classifiers for the Classification of PD Patients by the Odor Profile.** Table 3 shows the classification effect of the models. We used 80% for the training dataset and 20% for the validation dataset. It can be seen that the odor profile diagnosis model established by AB obtained the highest accuracy with 0.855 scores. Except for the model established by KNN, the models established by other classifiers were all greater than 0.800. The best F1 was 0.846 from AB. The RF model had the highest recall with a score of 0.981. The SVM model got the highest precision with a score of 0.882.

The relationship between specificity and sensitivity was obtained by plotting the ROC curve. As shown in Figure 5A, the area under the ROC curve indicated that the model established by the AIO system through SVM had good accuracy.

2.6. Medical Diagnostic Tests for the Established Model. Considering that the odor diagnosis model of PD had a better classification effect, the clinical application of the model was evaluated. To ensure the cleanliness of the test dataset, the researchers were blinded in this trial, and all participants' information was kept in the third party—the



**Figure 4.** ROC curve of the development cohort and validation cohort based on significant features: *X*-axis: false positive rate, *Y*-axis: true positive rate, purple line: development cohort, and pink line: validation cohort. The AUCs were 0.754 and 0.646, respectively.

Table 3. Model Evaluation Parameters

	evaluation parameters			
model	accuracy	recall	precision	F1
SVM	0.812	0.865	0.882	0.813
RF	0.841	0.981	0.836	0.818
KNN	0.754	0.980	0.754	0.689
AB	0.855	0.960	0.857	0.846
NB	0.841	0.924	0.875	0.835

Fourth Affiliated Hospital of Zhejiang University School of Medicine. After the analysis and model diagnosis experiment, the doctor will then send the participants' information to the experimenter for statistical analysis to evaluate the clinical application value of the model. The results of the evaluation of the clinical application for the five models are shown in Table 4. The ROC curve of the results is shown in Figure 5B.

Table 4 shows the results of the evaluation of the clinical application. We blindly diagnosed 12 PD and 12 HC. The odor diagnosis model constructed by RF obtained the highest accuracy with 0.792. We could find that the ensemble-learning

 Table 4. Results of the Evaluation of Clinical Application for

 Five Different Classifiers

+PV (%)	-PV (%)
66.7	66.7
73.3	88.9
52.4	66.7
71.4	80.0
80.0	57.9
	+PV (%) 66.7 73.3 52.4 71.4 80.0

classifiers—RF and AB—have higher accuracy than others. The highest sensitivity was 0.917 from RF and KNN. NB achieved the highest specificity with 0.916. The highest +PV, which doctors care about in clinical situations, was 0.800 from NB, and the highest –PV was 0.889 from RF. NB had the highest accuracy rate of TP in predicting positives. RF had the highest accuracy rate of TN in predicting negatives. However, the classification effect of SVM in the classifier did not reflect good results. The ROC curve in Figure 5B also showed that the ensemble classifiers had better results, and these classifiers also had the evaluation of the clinical application.

# 3. DISCUSSION

This research proposed fast GC combined with a SAW sensor with embedded ML algorithms to build an AIO system to diagnose PD through smell. ML was used to classify the sebaceous skin gas of PD patients based on the peaks in the chromatogram. There were three significant biomarkers (octanal, hexyl acetate, and perillic aldehyde) between PD patients and the control group. Using the three VOC biomarkers and the odor profile collected by the AIO system, the accuracies of classification between PD and HC were 70.8 and 79.2%, respectively.

The three significant biomarkers might be caused by different metabolic ways of PD patients. Hexyl acetate was usually found in many fruits and alcoholic beverages.<sup>36,37</sup> Perillic aldehyde was used in perfumes, cosmetics, and food.<sup>26</sup> The concentrations of hexyl acetate in HC were lower than those of PD, but the concentrations of perillic aldehyde in HC were higher than those in PD. It could be speculated that PD had a special metabolic ability for these two lipid hydrophobic metabolites (hexyl acetate and perillic aldehyde). SD was a



**Figure 5.** ROC curve analysis to evaluate the performance of different classifier construction models. Each colored line represents the ROC curve of the Parkinson's odor diagnosis model constructed by different classifiers: (A) development cohort: the ROC curve of the model. The AUCs of five classifiers (SVM, RF, KNN, AB, and NB) are 0.808, 0.868, 0.800, 0.929, and 0.914, respectively and (B) validation cohort: the ROC curve of the medical diagnostic tests. The AUCs of five classifiers are 0.681, 0.819, 0.729, 0.826, and 0.698, respectively.



**Figure 6.** (A) System design of the AIO system; solid red line: sampling mode system operating gas path; dotted blue line analyzing mode system operating gas path. (B) Process of clinical experiments: (1) basic information about the participants was recorded; (2) the gauze was placed on the participants' back to extract the skin sebum VOCs, and then, the sample gauze was placed in a glass bottle with an inert brown background gas; (3) the bottles were transported in ice packs; (4) the samples were taken back to the laboratory and placed in the refrigerator; and (5) analytical experiments were carried out on the collected samples by the AIO system to obtain odor profiles (created with BioRender.com).

typical nonmotor symptom of PD and was developed by the increasing sebum excretion and proliferation of *Malassezia* yeasts.<sup>38</sup> Increased sebum excretion in PD patients would increase yeast and the production of enzymes, leading to sebum inflammation.<sup>23</sup> Both neural and epidermal tissues originated from the ectoderm. Gioti described that the growth of *Malassezia* bacteria required specific exogenous lipids, which might be related to the increase of lipophilic molecules.<sup>39</sup> Both hexyl acetate and perillic aldehyde were lipophilic molecules and were insoluble in water, so they might be required for the growth of *Malassezia*, which might be related to PD. Hirayama pointed out that PD could cause abnormal sweat glands, leading to excessive sweating and night sweats.<sup>41,42</sup> Moreover, Agapiou

showed that octanal might be related to oxidative stress.<sup>43</sup> Besides, Puspita demonstrated that oxidative stress was one of the pathogeneses of PD.<sup>44</sup> Therefore, oxidative stress might cause an increase of the octanal content on the skin surface of PD patients. All potential explanations for changes in odor in PD indicated that the changes in skin physiology and skin metabolomics were highly specific to PD. This made them useful as biomarkers for identifying patients with PD. Besides, the clinical concentration range was expected to be 0.25–2.5 mM according to the experimental results from Trivedi's studies. The results from our study also supported their reports since the response of AIO was always ranging from 1 to 25 kHz in most of the chromatograms, which was referring to 0.25–2.5 mM of various markers by the calibration curves.

The results of the classification model based on the sebum odor profile are pretty good. The classification accuracies of PD and HC were 70.8 and 79.2%, respectively, using the significant biomarker features and odor profile. The model based on the odor profile has a highest sensitivity of 91.7%, a highest specificity of 91.6%, and a highest AUC of 0.826. It could be inferred that some peaks were not significant on the chromatogram and were also included in the algorithm when classifying based on the odor profile. These insignificant peaks also contributed to the classification, improving accuracy, sensitivity, and specificity. Besides, the smell profile, which covered many VOCs' peaks, also helped improve the sensitivity and specificity of the classification model based on the odor profile of sebum.

The fast, easy-to-use, and portable AIO system can balance sensitivity, specificity, linearity, dynamic range, detection limit, detection efficiency, and discrimination ability. Compared with the GC-MS, liquid chromatography-MS, or paper spray ionization coupled with ion mobility-MS, the AIO system greatly improves the detection speed and reduces the detection cost.<sup>26,27,30,45,46</sup> Compared with traditional clinical PD diagnosis methods, the AIO system is a fast and noninvasive method. It can also be widely used in hospitals, clinics, and homes as a PD patient screening or family self-health check method. PD is a chronic neurodegenerative disease caused by the loss of dopaminergic neurons in the substantia nigra. Therefore, when PD patients develop motor symptoms, they have already lost dopaminergic neurons.<sup>47</sup> The neurodegeneration process may be too fast, so it is essential to identify PD before widespread neuron loss occurs, and the AIO provides a possible solution.

However, there were several limitations to the present study. First, AIO used fast GC to separate mixed VOCs. The GC method has a limitation. According to the principle of GC separation, each peak represented a pure chemical compound that had a unique retention time to be recognized. However, in some circumstances, two or more compounds had quite close retention times, which would cause overlaps of the peaks in the fast GC separation. We used these retention times to identify the biomarkers (Table S2). However, it is hard to avoid the situation when the interfering compounds' retention time was the same as that of the biomarkers. Second, the diagnostic accuracy based on the classification largely depended on the size of the training set and the representativeness of the sample population. In our study population, the data distribution of samples generally needed to be relatively balanced, so the results obtained by the model may be expected to have high classification accuracy. Furthermore, the controlled equalization of samples in the PD and HC did not represent the distribution of the PD disease in real clinical settings, leading to the limited utility of the model. Third, the metabolic mechanism of hexyl acetate and perillic aldehyde in the organism was not clear. Moreover, the differentiation of dodecane, which was proved to be related to the UPDRS3, was not observed in this study, indicating the existence of interfering factors such as race. Nonetheless, this study provided some ideas worthy of further study. If more research supported our hypothesis, additional biomarkers such as hexyl acetate, perillic aldehyde, octanal, and dodecane could be incorporated into the risk score model to identify individuals at a high risk of PD. From a future perspective, when the high sensitivity and specificity of the AIO system is reproduced within more extensive studies, the AIO system might assist

clinicians in monitoring the extent of PD in patients who have PD or who may be at a high risk of PD.

# 4. CONCLUSIONS

In conclusion, we proposed a fast GC system combined with the SAW sensor with embedded ML algorithms to establish an AIO system for the diagnosis of PD through smell. This method presents a new possibility for the early diagnosis of PD. Compared with olfactory testing, sleep testing, and other solutions, the combination of the AIO system and ML may produce a new method of gaseous-assisted diagnosis of PD with an improved detection speed and a reduced detection cost. Moreover, the AIO system is a fast, easy-to-use, and noninvasive method that can be widely used in hospitals, clinics, and homes to screen, diagnose, and monitor PD treatment.

### 5. MATERIALS AND METHODS

5.1. Artificial Intelligence Olfactory System. 5.1.1. Design of the Fast GC System. The AIO system is a combination of GC and SAW sensor. The system comprises three modules: gas injection and preconcentration module, chromatographic separation module (GC), and sensor detection module (SAW). As shown in Figure 6A, an adsorbent tube filled with 10 mg of Tenax TA (60/80 mesh, purchased from Analytical Columns, Croydon, England), a six-way valve, and a vacuum pump was used to preconcentrate the sample gases for the injection. A 1 m long DB-1 capillary column (cut from an Agilent column of 10 m  $\times$  0.1 mm  $\times$  0.33 mm) with a direct resistive heating component (manufactured by Ningbo Oulaike Metal Capillary Technology Co., Ltd., Ningbo, China) was used to fast separate the compounds in the sample gases. A 36° Y-X cut quartz substrate double-ended resonant Rayleigh wave gas sensor with a center frequency of 500 MHz (manufactured by Hua Ying Electronics, Wukang, China) was used to detect the separated compounds with a sensitivity of -69,766 Hz/ng to mass deposition. Also, the AIO system uses Bluetooth to connect with the host computer.

The AIO system has two working states controlled by a sixway valve, including the sampling mode and the analyzing mode. The six-way valve was the key component of the AIO system. Under the normal conditions, the six-way valve was in the sampling mode, as indicated by the solid red line in Figure 6A. Under the control of the vacuum pump, the gas sample flowed from the gas inlet to an adsorbent tube, before the sample gets adsorbed into this tube. After adsorption, the sixway valve was switched to the analyzing mode, as indicated by the dotted blue line in Figure 6A. By instantaneously heating the adsorption tube, the analyte in the adsorbent was desorbed before flowing into the DB-1 capillary column for future separation. Additionally, the separated substances were tested for quality on the SAW.

5.1.2. Design of the Surface Acoustic Wave Sensor. A  $36^{\circ}$  Y-X cut quartz substrate double-ended resonant Rayleigh wave gas sensor with a center frequency of 500 MHz was used to detect the mass loading of the compounds separated by GC. The specific parameters of the SAW sensor are shown in Table 5. According to the mass deposition effect formula of the SAW device, the sensitivity of the sensor mass deposition can be calculated as -69,766 Hz/ng. The SAW sensor vibration spectrum measured by the spectrum analyzer is shown in

# Table 5. Parameters of the SAW Sensor

index	parameter
substrate materials	36° Y–X quartz
electrode material	aluminum
electrode thickness	200 nm
central frequency	500 MHz
input/output transducers	50.5 pairs
reflectors	350 in each side
transducer aperture	800 µm
input/output transducer cycle ( $\lambda$ )	6.3 μm
reflector cycle	λ
reflector and transducer spacing	λ
input/output transducer spacing	1.25λ

Figure S2. Also, the image of the SAW sensor is shown in Figure S1.

5.1.3. Operation Method of the AIO System. The initial temperatures of the injection port, six-way valve, capillary column, and sensor were, respectively, set to 80, 130, 30, and 35 °C. The flow rate of carrier helium gas was fixed at 1 mL/min. After preheating, the gas path was set to the sample mode. The sampling time was 20 s, and the analysis time was 30 s. Then, about 20 mL of sample gas was adsorbed into the Tenax TA adsorption tube. For gas analysis, the Tenax TA adsorption tube was heated. Next, the valve was used to change the gas path to the analysis mode. The initial temperature of the capillary column was maintained at 30 °C for 1 s and then increased to 120 °C at a rate of 6 °C/s. Finally, the SAW sensor was heated to 105 °C for cleaning. All the procedures, including sampling, separation, detection, and cooling, took 90 s in one analysis cycle. Moreover, the ambient temperature in the laboratory was controlled at about 22 °C.

**5.2.** Calibrating AIO by Standard Samples. *5.2.1. Reagents and Materials.* Octanal was ( $C_8H_{16}O$ , 98.0%) purchased from TOKYO Chemical Industry. Perillic aldehyde ( $C_{10}H_{14}O$ , 95.0%) was purchased from Shanghai Yuanye Biological Technology Co., Ltd. Hexyl acetate ( $C_8H_{16}O$ , 98.0%), nonane ( $C_9H_{20}$ , 98.0%), decane ( $C_{10}H_{22}$ , 99.0%), undecane ( $C_{11}H_{24}$ , 99.5%), dodecane ( $C_{12}H_{26}$ , 98.0%), and tetradecane ( $C_{14}H_{30}$ , 98.0%) were obtained from Aladdin Industrial Corporation.

Tridecane ( $C_{13}H_{28}$ , 98.0%) and pentadecane ( $C_{15}H_{32}$ , 98.0%) were purchased from Macklin. All solutions were stored at 4 °C and protected from light.

5.2.2. Preparation of Standard Samples. The first part is to test the AIO system's ability to distinguish the mixed substances. Standard solutions of nonane, decane, undecane, dodecane, tridecane, tetradecane, and pentadecane were prepared into a mixed solution, and all solutions were diluted to parts per million of the original concentration. The second part is the calibration of the linear repeatability of the AIO system. Standard samples of four biomarkers (octanal, hexyl acetate, perillic aldehyde, and dodecane) were applied to calibrate AIO. Four biomarkers' reagents were used to prepare standard solutions, and the corresponding solvents were used to configure nine different types: 0.025, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 25, and 50 mM concentrations of four marker solutions. Among them, the solvent of octanal, perillic aldehyde, and dodecane reagents was ethanol, and the solvent of the hexyl acetate reagent was methanol. Then, AIO was calibrated by injecting samples (0.02  $\mu$ L) with a microinjector (1  $\mu$ L, manufactured by Shanghai High Pigeon Industry and Trade Co. Ltd.), each concentration was measured 3 times, and linear regression analysis was performed (Section 5.4).

**5.3. Clinical Experiment Design.** The experiments were conducted in the Fourth Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China. This study was approved by the ethics committee of the Fourth Affiliated Hospital of Zhejiang University School of Medicine (approval no. K2020052, 15 June 2020). A total of 87 participants were recruited in our experiment, including 43 PD patients and 44 HC. The patients in the study were diagnosed as PD and were staged by the neurologist according to the H–Y clinical staging scale in the hospital. The early stage includes stages I and II, and the middle and late stages are stages III–V. Basic clinical data (patients' ID, age, weight, drug exposure history, years of education, high-protein diet, smoking and drinking history, cerebrovascular history, disease duration, and skin condition) were collected.

Each participant had signed informed consent. To ensure the accuracy of the results, subjects are required to comply with the following requirements: (1) no bathing, body lotion, or other cosmetics are allowed 12 h before the sample collection; (2) subjects are required to refrain from exercise for 12 h before sample collection; (3) no perfume shall be sprayed 12 h before sample collection; and (4) fasting 4 h before collection.

Initially, considering that the subjects are mostly elderly, paper questionnaires were used to record relevant information of the participants, and the information was organized and stored on the computer. Second, each participant was swabbed using a medical gauze ( $7.5 \times 7.5$  cm) on the upper back to collect sebum samples. The gauze with the sebum sample was sealed in background-inert plastic bags and transported from the hospital to the laboratory using the ice bag. The samples were stored in a -25 °C refrigerator. Next, the above method (Section 5.1) was used to set up the AIO system for the chromatographic analysis of the experiment. At last, the data obtained by the analysis were displayed on the host computer via Bluetooth. The process of the clinical experiments is shown in Figure 6B.

**5.4. Data Analysis.** VOC data and data from each volunteer were recorded in a dataset. Statistical analysis was done by Statistical Package for the Social Sciences (PASW Statistics 18, IBM Corp., Armonk, NY, USA), GraphPad (Prism 5, GraphPad Software, Inc., La Jolla, CA, USA), and Python (Python 3.7, the MathWorks, Inc., Natick, MA, USA) for Windows.

The linear regression was used to fit the average sensor response to the concentrations of the three calibration reagent solutions (octanal, hexyl acetate, and perillic aldehyde). The goodness of fit was evaluated using the correlation R. The Mann–Whitney U test and Kolmogorov–Smirnov Z test were used to select the significant features. The differences with a p-value less than 0.05 were considered to be statistically significant. Python was used to create the five different classifiers (SVM, RF, KNN, AB, and NB). Parameters of the classifiers were selected based on GridsearchCV, and they were as follows. The linear kernel was utilized for SVM. The number of trees was 25 for RF. The k parameter was 10 for KNN for all experiments. The area under the ROC curve (AUC) was a performance indicator for binary classification problems based on classification.

# ASSOCIATED CONTENT

#### **Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c05060.

Results of the nonparametric test for the four biomarkers; retention time of four standard biomarkers by the AIO system; SAW sensor vibration frequency spectrum; and physical image of the SAW sensor (PDF)

# AUTHOR INFORMATION

# **Corresponding Authors**

Jun Liu – Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China; Email: liujun@zju.edu.cn

Xing Chen – Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China; Present Address: Zhejiang Sir Run Run Shaw Hospital, Department of Medicine, Zhejiang University, Hangzhou, Zhejiang 310027, China; orcid.org/0000-0002-9201-4906; Email: cnxingchen@ zju.edu.cn

#### Authors

- Wei Fu Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China; ⊙ orcid.org/0000-0003-4581-8909
- Linxin Xu Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China
- **Qiwen Yu** Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China
- Jiajia Fang Department of Neurology, the Fourth Affiliated Hospital, Zhejiang University School of Medicine, Yiwu City, Zhejiang Province 322000, P. R. China
- Guohua Zhao Department of Neurology, the Fourth Affiliated Hospital, Zhejiang University School of Medicine, Yiwu City, Zhejiang Province 322000, P. R. China
- Yi Li Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China
- **Chenying Pan** Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China
- Hao Dong Research Center for Intelligent Sensing, Zhejiang Lab, Hangzhou 311100, China
- Di Wang Research Center for Intelligent Sensing, Zhejiang Lab, Hangzhou 311100, China; Orcid.org/0000-0003-1581-4982
- Haiyan Ren Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China
- Yi Guo Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China
- **Qingjun Liu** Department of Biomedical Engineering, Key Laboratory of Biomedical Engineering of Ministry of

Education of China, Zhejiang University, Hangzhou, Zhejiang 310027, China; orcid.org/0000-0002-9442-9355

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c05060

# Notes

The authors declare no competing financial interest.

# ACKNOWLEDGMENTS

We thank participating volunteers for their effort and time. Funding supports: National Natural Science Foundation of China (no. 82172064), National Key Research and Development Program of China (no. 2019YFC1711800), Zhejiang Public Welfare Technology Application Research Project (no. LGF20H090011), Key Research and Development Program of Shaanxi (no. 2020ZDLSF04-03), Major Scientific Project of Zhejiang Lab (no. 2020MC0AD01), Zhejiang Provincial Natural Science Foundation of China (no. LQ21F010003), a China Postdoctoral Science Foundation (no. 2020M681952), and Major Consulting Project of the Chinese Academy of Engineering (no. 2019-ZD-6-03).

### REFERENCES

(1) Poewe, W.; Seppi, K.; Tanner, C. M.; Halliday, G. M.; Brundin, P.; Volkmann, J.; Schrag, A.-E.; Lang, A. E. Parkinson disease. *Nat. Rev. Dis. Primers* **2017**, *3*, 17013.

(2) Savica, R.; Rocca, W. A.; Ahlskog, J. E. When does Parkinson disease start? *Arch. Neurol.* 2010, 67, 798-801.

(3) Balestrino, R.; Schapira, A. H. V. Parkinson disease. Eur. J. Neurol. 2020, 27, 27–42.

(4) von Campenhausen, S.; Bornschein, B.; Wick, R.; Bötzel, K.; Sampaio, C.; Poewe, W.; Oertel, W.; Siebert, U.; Berger, K.; Dodel, R. Prevalence and incidence of Parkinson's disease in Europe. *Eur. Neuropsychopharmacol.* **2005**, *15*, 473–490.

(5) Marras, C.; Beck, J. C.; Bower, J. H.; Roberts, E.; Ritz, B.; Ross, G. W.; Abbott, R. D.; Savica, R.; Van Den Eeden, S. K.; Willis, A. W.; Tanner, C. M. Prevalence of Parkinson's disease across North America. *npj Parkinson's Dis.* **2018**, *4*, 21.

(6) Muangpaisan, W.; Hori, H.; Brayne, C. Systematic review of the prevalence and incidence of Parkinson's disease in Asia. *J. Epidemiol.* **2009**, *19*, 281–293.

(7) Rocca, W. A. The burden of Parkinson's disease: a worldwide perspective. *Lancet Neurol.* **2018**, *17*, 928–929.

(8) GBD 2016 Neurology Collaborators. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **2018**, *17*, 939–953.

(9) McKinley, J. E.; Perkins, A. Neurologic Conditions: Parkinson Disease. *FP Essent.* **2019**, 477, 16–21.

(10) Newaj, R. Parkinson's: an overview. *Frontshop* **2017**, *2017*, 38. (11) Kobylecki, C. Update on the diagnosis and management of Parkinson's disease. *Clin. Med.* **2020**, *20*, 393–398.

(12) Hoehn, M. M.; Yahr, M. D. Parkinsonism: onset, progression, and mortality. 1967. *Neurology* **1998**, *50*, 318.

(13) Goetz, C. G.; Poewe, W.; Rascol, O.; Sampaio, C.; Stebbins, G. T.; Counsell, C.; Giladi, N.; Holloway, R. G.; Moore, C. G.; Wenning, G. K.; Yahr, M. D.; Seidl, L. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov. Disord.* **2004**, *19*, 1020–1028.

(14) Goetz, C. G.; Tilley, B. C.; Shaftman, S. R.; Stebbins, G. T.; Fahn, S.; Martinez-Martin, P.; Poewe, W.; Sampaio, C.; Stern, M. B.; Dodel, R.; Dubois, B.; Holloway, R.; Jankovic, J.; Kulisevsky, J.; Lang, A. E.; Lees, A.; Leurgans, S.; LeWitt, P. A.; Nyenhuis, D.; Olanow, C. W.; Rascol, O.; Schrag, A.; Teresi, J. A.; van Hilten, J. J.; LaPelle, N. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* 2008, 23, 2129–2170.

(15) Martinez-Martin, P.; Rodriguez-Blazquez, C.; Alvarez-Sanchez, M.; Arakaki, T.; Bergareche-Yarza, A.; Chade, A.; Garretto, N.; Gershanik, O.; Kurtis, M. M.; Martinez-Castrillo, J. C.; Mendoza-Rodriguez, A.; Moore, H. P.; Rodriguez-Violante, M.; Singer, C.; Tilley, B. C.; Huang, J.; Stebbins, G. T.; Goetz, C. G. Expanded and independent validation of the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS). J. Neurol. 2013, 260, 228–236.

(16) Storch, A.; Schneider, C. B.; Klingelhöfer, L.; Odin, P.; Fuchs, G.; Jost, W. H.; Martinez-Martin, P.; Koch, R.; Reichmann, H.; Chaudhuri, K. R.; Ebersbach, G. Quantitative assessment of non-motor fluctuations in Parkinson's disease using the Non-Motor Symptoms Scale (NMSS). *J. Neural Transm.* **2015**, *122*, 1673–1684.

(17) Martinez-Martin, P.; Rodriguez-Blazquez, C.; Abe, K.; Bhattacharyya, K. B.; Bloem, B. R.; Carod-Artal, F. J.; Prakash, R.; Esselink, R. A. J.; Falup-Pecurariu, C.; Gallardo, M.; Mir, P.; Naidu, Y.; Nicoletti, A.; Sethi, K.; Tsuboi, Y.; van Hilten, J. J.; Visser, M.; Zappia, M.; Chaudhuri, K. R. International study on the psychometric attributes of the non-motor symptoms scale in Parkinson disease. *Neurology* **2009**, *73*, 1584–1591.

(18) Kulisevsky, J.; Pagonabarraga, J. Cognitive impairment in Parkinson's disease: tools for diagnosis and assessment. *Mov. Disord.* **2009**, *24*, 1103–1110.

(19) Pagonabarraga, J.; Kulisevsky, J.; Llebaria, G.; García-Sánchez, C.; Pascual-Sedano, B.; Gironell, A. Parkinson's disease-cognitive rating scale: a new cognitive scale specific for Parkinson's disease. *Mov. Disord.* **2008**, *23*, 998–1005.

(20) Pagano, G.; Niccolini, F.; Politis, M. Imaging in Parkinson's disease. *Clin. Med.* **2016**, *16*, 371–375.

(21) de la Fuente-Fernández, R. Role of DaTSCAN and clinical diagnosis in Parkinson disease. *Neurology* **2012**, *78*, 696–701.

(22) Parnetti, L.; Gaetani, L.; Eusebi, P.; Paciotti, S.; Hansson, O.; El-Agnaf, O.; Mollenhauer, B.; Blennow, K.; Calabresi, P. CSF and blood biomarkers for Parkinson's disease. *Lancet Neurol.* **2019**, *18*, 573–586.

(23) Nazik, H.; Yildiz, B. T. Evaluation of skin disorders, skin sebum and moisture in patients with Parkinson's disease. *Neurol. Asia* **2019**, *24*, 249–254.

(24) Mastrolonardo, M.; Diaferio, A.; Logroscino, G. Seborrheic dermatitis, increased sebum excretion, and Parkinson's disease: a survey of (im)possible links. *Med. Hypotheses* **2003**, *60*, 907–911.

(25) Ravn, A.-H.; Thyssen, J. P.; Egeberg, A. Skin disorders in Parkinson's disease: potential biomarkers and risk factors. *Clin., Cosmet. Invest. Dermatol.* **2017**, *10*, 87–92.

(26) Trivedi, D. K.; Sinclair, E.; Xu, Y.; Sarkar, D.; Walton-Doyle, C.; Liscio, C.; Banks, P.; Milne, J.; Silverdale, M.; Kunath, T.; Goodacre, R.; Barran, P. Discovery of Volatile Biomarkers of Parkinson's Disease from Sebum. *ACS Cent. Sci.* **2019**, *5*, 599–606.

(27) Sinclair, E.; Walton-Doyle, C.; Sarkar, D.; Hollywood, K. A.; Milne, J.; Lim, S. H.; Kunath, T.; Rijs, A. M.; de Bie, R. M. A.; Silverdale, M.; Trivedi, D. K.; Barran, P. Validating Differential Volatilome Profiles in Parkinson's Disease. *ACS Cent. Sci.* **2021**, *7*, 300–306.

(28) Sinclair, E.; Trivedi, D. K.; Sarkar, D.; Walton-Doyle, C.; Milne, J.; Kunath, T.; Rijs, A. M.; de Bie, R. M. A.; Goodacre, R.; Silverdale, M.; Barran, P. Metabolomics of sebum reveals lipid dysregulation in Parkinson's disease. *Nat. Commun.* **2021**, *12*, 1592.

(29) Pereira, P. A. B.; Trivedi, D. K.; Silverman, J.; Duru, I. C.; Paulin, L.; Auvinen, P.; Scheperjans, F. Multiomics implicate gut microbiota in altered lipid and energy metabolism in Parkinson's disease. *medRxiv* **2021**, 2021, 21258035.

(30) Tsuda, T.; Nonome, T.; Goto, S.; Takeda, J.-i.; Tsunoda, M.; Hirayama, M.; Ohno, K. Application of Skin Gas GC/MS Analysis for Prediction of the Severity Scale of Parkinson's Disease. *Chromatography* **2019**, *40*, 149–155.

(31) Mebazaa, R.; Rega, B.; Camel, V. Analysis of human male armpit sweat after fenugreek ingestion: Characterization of odour

active compounds by gas chromatography coupled to mass spectrometry and olfactometry. *Food Chem.* **2011**, *128*, 227–235.

(32) Bernier, U. R.; Kline, D. L.; Barnard, D. R.; Schreck, C. E.; Yost, R. A. Analysis of human skin emanations by gas chromatography/ mass spectrometry. 2. Identification of volatile compounds that are candidate attractants for the yellow fever mosquito (Aedes aegypti). *Anal. Chem.* **2000**, *72*, 747–756.

(33) Staples, E. J.; Viswanathan, S. Odor Detection and Analysis Using GC/SAW zNose; Electronic Sensor Technology, School of Engineering and Technology, National University, 2015.

(34) He, S.; Liu, J.; Liu, M. The SAW gas chromatograph and its applications in the public security. 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing; IEEE Computer Society, 2013; pp 1710–1713.

(35) Li, S. Recent Developments in Human Odor Detection Technologies. J. Forensic Sci. Criminol. 2014, 1, 1–2.

(36) Dennis, E. G.; Keyzers, R. A.; Kalua, C. M.; Maffei, S. M.; Nicholson, E. L.; Boss, P. K. Grape contribution to wine aroma: production of hexyl acetate, octyl acetate, and benzyl acetate during yeast fermentation is dependent upon precursors in the must. *J. Agric. Food Chem.* **2012**, *60*, 2638–2646.

(37) Gómez, E.; Ledbetter, C. A. Comparative study of the aromatic profiles of two different plum species: Prunus salicina Lindl and Prunus simonii L. J. Sci. Food Agric. **1994**, 65, 111–115.

(38) Gupta, A. K.; Madzia, S. E.; Batra, R. Etiology and management of Seborrheic dermatitis. *Dermatology* **2004**, *208*, 89–93.

(39) Gioti, A.; Nystedt, B.; Li, W.; Xu, J.; Andersson, A.; Averette, A. F.; Münch, K.; Wang, X.; Kappauf, C.; Kingsbury, J. M.; Kraak, B.; Walker, L. A.; Johansson, H. J.; Holm, T.; Lehtiö, J.; Stajich, J. E.; Mieczkowski, P.; Kahmann, R.; Kennell, J. C.; Cardenas, M. E.; Lundeberg, J.; Saunders, C. W.; Boekhout, T.; Dawson, T. L.; Munro, C. A.; de Groot, P. W.; Butler, G.; Heitman, J.; Scheynius, A. Genomic insights into the atopic eczema-associated skin commensal yeast Malassezia sympodialis. *mBio* **2013**, *4*, e00572–00512.

(40) Hirayama, M. Sweating dysfunctions in Parkinson's disease. J. Neurol. 2006, 253, Vii42–Vii47.

(41) Duffy, E.; Albero, G.; Morrin, A. Headspace Solid-Phase Microextraction Gas Chromatography-Mass Spectrometry Analysis of Scent Profiles from Human Skin. *Cosmetics* **2018**, *5*, 62.

(42) Dormont, L.; Bessière, J.-M.; Cohuet, A. Human skin volatiles: a review. J. Chem. Ecol. 2013, 39, 569-578.

(43) Agapiou, A.; Amann, A.; Mochalski, P.; Statheropoulos, M.; Thomas, C. Trace detection of endogenous human volatile organic compounds for search, rescue and emergency applications. *TrAC, Trends Anal. Chem.* **2015**, *66*, 158–175.

(44) Puspita, L.; Chung, S. Y.; Shim, J.-w. Oxidative stress and cellular pathologies in Parkinson's disease. *Mol. Brain* **2017**, *10*, 53.

(45) Sarkar, D.; Trivedi, D.; Sinclair, E.; Lim, S.; Walton-Doyle, C.; Jafri, K.; Milne, J.; Silverdale, M.; Barran, P. Rapid Diagnosis of Parkinson's Disease from Sebum using Paper Spray Ionisation Ion Mobility Mass Spectrometry. *ChemRxiv* 2020, DOI: 10.26434/ chemrxiv.12517385.v1.

(46) Sinclair, E.; Trivedi, D.; Sarkar, D.; Walton-Doyle, C.; Milne, J.; Kunath, T.; Rijs, A.; Debie, R.; Goodacre, R.; Silverdale, M.; Barran, P. Sebum: A Window into Dysregulation of Mitochondrial Metabolism in Parkinson's Disease. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.11603613.v2.

(47) Niswender, C. M.; Johnson, K. A.; Weaver, C. D.; Jones, C. K.; Xiang, Z.; Luo, Q.; Rodriguez, A. L.; Marlo, J. E.; de Paulis, T.; Thompson, A. D.; Days, E. L.; Nalywajko, T.; Austin, C. A.; Williams, M. B.; Ayala, J. E.; Williams, R.; Lindsley, C. W.; Conn, P. J. Discovery, characterization, and antiparkinsonian effect of novel positive allosteric modulators of metabotropic glutamate receptor 4. *Mol. Pharmacol.* **2008**, *74*, 1345–1358.