

Diagnosis and Classification of 17 Diseases from 1404 Subjects via Pattern Analysis of Exhaled Molecules

Morad K. Nakhleh,[†] Haitham Amal,[†] Raneen Jeries,[†] Yoav Y. Broza,[†] Manal Aboud,[†] Alaa Gharra,[†] Hodaya Ivgi,[†] Salam Khatib,[†] Shifaa Badarneh,[†] Lior Har-Shai,[‡] Lea Glass-Marmor,[‡] Izabella Lejbkowicz,[‡] Ariel Miller,[‡] Samih Badarny,[§] Raz Winer,[§] John Finberg,^{||} Sylvia Cohen-Kaminsky,^{||} Frédéric Perros,^{||} David Montani,^{||} Barbara Girerd,^{||} Gilles Garcia,^{||} Gérald Simonneau,^{||} Farid Nakhoul,[#] Shira Baram,[¶] Raed Salim,[¶] Marwan Hakim,[□] Maayan Gruber,[□] Ohad Ronen,[□] Tal Marshak,[□] Ilana Doweck,[□] Ofer Nativ,[○] Zaher Bahouth,[○] Da-you Shi,[●] Wei Zhang,[●] Qing-ling Hua,[●] Yue-yin Pan,[●] Li Tao,[●] Hu Liu,[●] Amir Karban,[△] Eduard Koifman,[△] Tova Rainis,[△] Roberts Skapars,[△] Armands Sivins,[△] Guntis Ancans,[△] Inta Liepniece-Karele,[△] Ilze Kikuste,[△] Ieva Lasina,[△] Ivars Tolmanis,[△] Douglas Johnson,[○] Stuart Z. Millstone,[●] Jennifer Fulton,^{††} John W. Wells,[●] Larry H. Wilf,^{††} Marc Humbert,^{||} Marcis Leja,[†] Nir Peled,^{§§} and Hossam Haick*,[†] and Hossam Haick*

[†]Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion—Israel Institute of Technology, Haifa 3200003, Israel

[‡]Division of Neuroimmunology and Multiple Sclerosis Center, Carmel Medical Center and Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 31096, Israel

[§]Movement Disorders Clinic, Department of Neurology, Carmel Medical Center, and Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 31096, Israel

^{||}Department of Molecular Pharmacology, Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 31096, Israel

[¶]Univ. Paris-Sud, Faculté de Médecine, Université Paris-Saclay, AP-HP, Centre National de Référence de l'Hypertension Pulmonaire Sévère, Département Hospitalo-Universitaire (DHU) Thorax Innovation, Service de Pneumologie, Hôpital de Bicêtre, UMRS_999, INSERM and Univ. Paris-Sud, Laboratoire d'Excellence (LabEx) en Recherche sur le Médicament et l'Innovation Thérapeutique (LERMIT), Centre Chirurgical Marie Lannelongue, Le Plessis Robinson 92350, France

[#]Department of Nephrology and Hypertension Baruch Padeh Medical Center, Poriya 15208, Israel

[○]Department of Obstetrics and Gynecology, Emek Medical Center, Afula 18101, and Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 31096, Israel

[□]Department of Obstetrics and Gynecology, Nazareth Hospital EMMS, Nazareth, and Faculty of Medicine in the Galilee, Bar Ilan University, Ramat Gan, Israel

[■]The Department of Otolaryngology Head and Neck Surgery, Carmel Medical Center, Haifa 3436212, Israel

[○]Department of Urology, Bnai Zion Medical Center, Haifa 31048, Israel

[●]Department of Oncology, The First Affiliated Hospital of Anhui Medical University, Hefei 230032, China

[△]Internal Medicine C and Gastroenterology Departments, Rambam Medical Center, Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 3525408, Israel

[▲]Department of Gastroenterology, Bnai Zion Hospital and Rappaport Family Faculty of Medicine, Technion—Israel Institute of Technology, Haifa 31096, Israel

[▽]Faculty of Medicine, University of Latvia, Digestive Diseases, Riga East University Hospital, 19 Rainisboulev, LV1586 Riga, Latvia

[▼]Digestive Diseases Centre, GASTRO, 6 Linezeraiela, LV1006 Riga, Latvia

[○]Department of Radiation Oncology, Baptist Cancer Institute (BCI), 1235 San Marco Boulevard, Suite100, Jacksonville, Florida 32207, United States

[●]Pulmonary and Critical Care Associates, Orange Park, Florida 32073, United States

Received: July 24, 2016

Accepted: December 2, 2016

Published: December 21, 2016

^{††}Pulmonary Diseases, Baptist Medical Center, Jacksonville, Florida 32217, United States

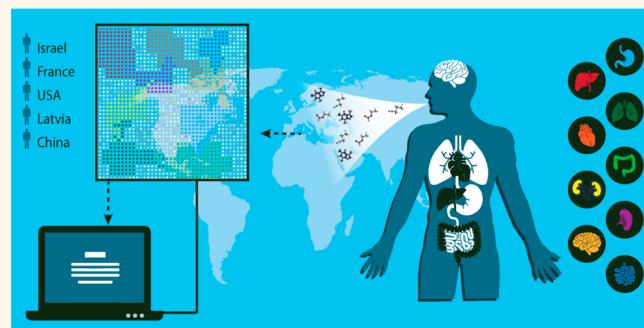
^{‡‡}Oncologic Imaging Division, Florida Radiation Oncology Group, Jacksonville, Florida 32217, United States

^{§§}Thoracic Cancer Unit, Davidoff Cancer Center, RMC, Kaplan Street, Petach Tiqwa 49100, Israel

Supporting Information

ABSTRACT: We report on an artificially intelligent nanoarray based on molecularly modified gold nanoparticles and a random network of single-walled carbon nanotubes for noninvasive diagnosis and classification of a number of diseases from exhaled breath. The performance of this artificially intelligent nanoarray was clinically assessed on breath samples collected from 1404 subjects having one of 17 different disease conditions included in the study or having no evidence of any disease (healthy controls). Blind experiments showed that 86% accuracy could be achieved with the artificially intelligent nanoarray, allowing both detection and discrimination between the different disease conditions examined. Analysis of the artificially intelligent nanoarray also showed that each disease has its own unique breathprint, and that the presence of one disease would not screen out others. Cluster analysis showed a reasonable classification power of diseases from the same categories. The effect of confounding clinical and environmental factors on the performance of the nanoarray did not significantly alter the obtained results. The diagnosis and classification power of the nanoarray was also validated by an independent analytical technique, *i.e.*, gas chromatography linked with mass spectrometry. This analysis found that 13 exhaled chemical species, called volatile organic compounds, are associated with certain diseases, and the composition of this assembly of volatile organic compounds differs from one disease to another. Overall, these findings could contribute to one of the most important criteria for successful health intervention in the modern era, *viz.* easy-to-use, inexpensive (affordable), and miniaturized tools that could also be used for personalized screening, diagnosis, and follow-up of a number of diseases, which can clearly be extended by further development.

KEYWORDS: sensor, nanoparticle, carbon nanotube, noninvasive, diagnosis, disease, breath, volatile organic compound



Since antiquity (~400 BC), physicians learned to evaluate their patients by exhaled volatile organic compounds (VOCs), among other means, that are linked to diseases.^{1,2} For example, the stools and urine of infant noblemen were smelt daily by their physicians. This approach rests on the fact that a compendium of VOCs with relatively low molecular weight expresses distinct and immediate changes as a consequence to pathophysiological processes occurring and altering the body's metabolism.^{2–8} Experimental findings show that VOCs can be detected directly from the headspace of the affected cells, blood, exhaled breath, and/or other body fluids.^{4,9} Of these body fluids, exhaled breath is the most accessible and useful VOC source for monitoring bodily health and disorder.^{4,9,10} This is because exhaled breath is easily and noninvasively accessible; it is suitable for high compliance; it provides a matrix of relative low complexity, and it can be handled safely and repeated as frequently as desired.

Detection of diseases from exhaled breath has been shown in different fields of medicine, particularly infectiology,^{11–13} respiratory medicine,^{14–19} and oncology.^{3,5,20–32} However, for exhaled breath analysis to move to the next stage of sophistication, by meeting the unmet needs of modern clinical practice, there is a dire need to extend the exhaled analysis approach so that it can not only facilitate disease diagnosis but also classify the disease condition. While disease diagnosis is the recognition of the presence of a specific disease, classification is extremely important for understanding the etiology, pathogenesis, and effective therapy of the patient's disease.^{33,34} Indeed, classification of disease is based on finding as many candidate diseases or conditions as possible that give the same signs or symptoms. This is followed by a process of elimination,

or at least of rendering the entries more or less probable by further medical tests and other processing, until one can reach the point where only one candidate disease or condition remains as the most probable.^{33,34} The final result may also provide a list of possible conditions, ranked in order of probability or severity.

Considering modern criteria in clinical practice, analysis of exhaled breath should be therefore low-cost, low-energy, ultra-miniaturized, easily repeated at specific time intervals, and have little or no impact on the day-to-day activity of the person diagnosed.³⁵ Nanotechnology-based chemical sensor matrices based on nanomaterials could serve as the translational tool leading from fundamental research to the modern point-of-care practice.^{6,7,22,36–38} This is because chemical sensor matrices are significantly smaller, easier to use, and less expensive than other sensing devices. Basically, recognition of VOCs by nanotechnology-based sensors can be achieved by selective detection of (pre)identified VOCs.^{7,36,39–41} This approach is useful to detect specific well-defined VOCs in the presence of interfering gaseous species or background using a highly selective receptor designed for this purpose, which is a laborious business.⁴² Despite advances in detection of VOCs by highly selective nanomaterial-based recognition methods, this has only been possible to date for indicating VOCs from a relatively narrow spectrum of diseases. While specific recognition in controlled backgrounds and interferences is achievable, most diseases today cannot be properly identified by individual VOCs alone.^{3,7,36,37} An additional limitation is the problem of synthesizing highly selective nanomaterials for each VOC, notably when they are nonpolar.⁴²

A complementary approach for disease detection and classification of a wider variety of diseases relies on cross-reactive

(*i.e.*, semiselective) nanotechnology-based sensor arrays, using pattern recognition;^{3,7,36–38} we refer to this approach as using an “artificially intelligent nanoarray”.^{43–46} In contrast to the selective method, an artificially intelligent nanoarray is more suitable for rapid diagnostic methods in which evaluation of a VOC compendium is qualitative and semiquantitative, with selectivity being achieved through pattern recognition of the compendium.^{3,7,36,37} Due to cross-reactivity, each sensor responds to a variety of VOCs, thereby allowing sensing and analysis of individual components from multicomponent samples.^{32,43,44,47,48} The concept of the artificially intelligent nanoarray is based on the ability of each sensor to detect all or part of the sample compounds. Although these sensors may have a sensitivity to a specific analyte (or VOC) lower than that of a selective sensor, they are more versatile in detecting multicomponent and complex VOC mixtures in different atmospheres (including those for which the (nano)arrays were not originally designed).^{32,43,44,47,48} Artificially intelligent nanoarrays of different composition were assessed in a series of separate laboratory (preclinical) and clinical studies for the detection of a wide range of cancerous and noncancerous diseases.^{8,14,23–32,43,45,47,49–57} In these studies, disease detection was mostly carried out with reference to healthy control groups, without examining correlated and uncorrelated clinical confounding factors. Clinical classification itself was beyond the focus of these studies.

In the present study, a multicenter clinical assessment was carried out to examine the probability of a single artificially intelligent nanoarray to detect and classify a range of disease types to explore the potential of this approach to modern clinical practice (Figure 1). A total of 2808 breath samples were collected from 1404 subjects having one of 17 different

disease types that had been put into three main categories (cancerous, inflammatory, and neurological diseases) and one group of uncorrelated diseases, as well as a group of healthy controls. The effects of primary demographic and environmental confounding factors (*e.g.*, sex, age, smoking habits, and geographical location) on the diagnosis and the classification results were examined. This was especially important because confounding factors are defined as factors that cause or prevent the diagnosis/classification of interest, are not intermediate variables, and are not associated with the diagnosis/classification under investigation. They give rise to situations in which the effects of the different processes are not separated, the contribution of the causal factors cannot be separated, or the measure of the effect of exposure (or risk) is distorted because of its association with other factors influencing the outcome. To gain an understanding of the outcome of the artificially intelligent nanoarrays, all breath samples were examined using an independent lab-based analytical technique referred to as gas chromatography linked with mass spectrometry (GC-MS). In contrast to an artificially intelligent nanoarray that analyzes collective breath VOC patterns, GC-MS identifies and quantifies specific components (VOCs) (Figure 1).

RESULTS

An artificially intelligent nanoarray that is based on chemiresistive layers of molecularly modified gold nanoparticles and random network of single-wall carbon nanotubes was designed and fabricated (detailed information can be found in the [Experimental Methods](#)). The inorganic nanomaterial-based species in these sensors (*i.e.*, gold nanoparticles or single-wall carbon nanotubes) provide the electrical conductivity, whereas the organic layer functions as a sensing layer (recognition element) for adsorbed VOCs.^{7,36,58} The sensing mechanism of

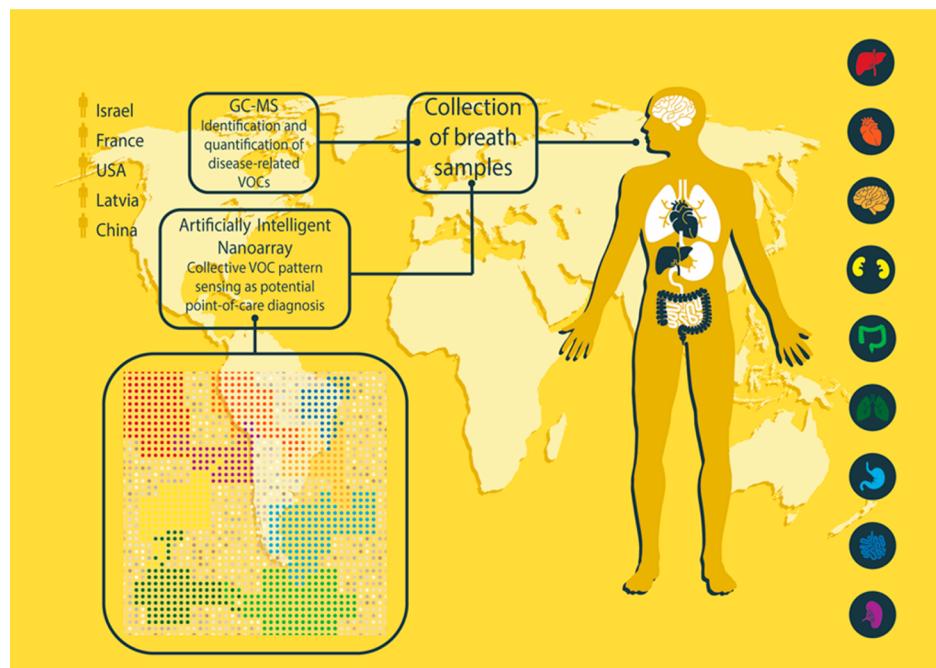


Figure 1. Schematic representation of the concept and design of the study. It involved collection of breath samples from 1404 subjects in 14 departments in nine clinical centers in five different countries (Israel, France, USA, Latvia, and China). The population included 591 healthy controls and 813 patients diagnosed with one of 17 different diseases: lung cancer, colorectal cancer, head and neck cancer, ovarian cancer, bladder cancer, prostate cancer, kidney cancer, gastric cancer, Crohn's disease, ulcerative colitis, irritable bowel syndrome, idiopathic Parkinson's, atypical Parkinsonism, multiple sclerosis, pulmonary arterial hypertension, pre-eclampsia, and chronic kidney disease. One breath sample obtained from each subject was analyzed with the artificially intelligent nanoarray for disease diagnosis and classification, and a second was analyzed with GC-MS for exploring its chemical composition.

these sensors can be affected by one or a combination of the following mechanisms. Sorption of VOCs into the organic film affects electron tunneling by reversible swelling or aggregation of the layers, which increases or decreases interparticle distance, leading, respectively, to an increase or decrease in the electrical resistance of the film.^{6,36,58} Another possible sensing mechanism that does not have steric changes within the sensing layer is charge transfer from or to the inorganic nanomaterial that develops on exposure to VOCs.^{6,36,58} In fact, the dielectric constant of the organic layer might change significantly when that of sorbed vapor differs significantly.^{6,36,58} The permittivity of the organic matrix surrounding the metal cores increases due to the higher dielectric constant of the sorbed vapor than that of the organic layer (*e.g.*, water, methanol). In these cases, tunneling activation energy decreases, leading to a decrease in the electrical resistance of the sensing film. Hence, sorbed vapor with a lower dielectric constant results in an increase in resistivity (*e.g.*, toluene, *n*-hexane).^{6,36,58} The chemical diversity of both conductive inorganic nanomaterials and organic layers results in the sensors responding differently to breath VOCs, which creates unique fingerprints in resistance changes. Selection of the conductive inorganic nanomaterials and organic layers can also be accurately tailored to a desired sensing application.

In conjunction with artificial intelligence methods, the nanoarray was used for a meta-analysis of several groups of subjects under real-world circumstances, each manifesting a specific health condition (Figure 1). This analysis was carried out on breath samples collected in a controlled manner from 1404 eligible subjects collected between January 2011 and June 2014 from 14 departments in nine clinical centers in five different countries (Israel, France, USA, Latvia, and China). Of the subjects, 813 were patients diagnosed with one of the following 17 diseases: chronic kidney failure (CKD), idiopathic Parkinson's disease (IPD), atypical Parkinsonism (PDISM), multiple sclerosis (MS), Crohn's disease (CD), ulcerative colitis (UC), irritable bowel syndrome (IBS), pulmonary arterial hypertension (PAH), pre-eclampsia in pregnant women (PET), head and neck cancer (HNC), lung cancer (LC), colorectal cancer

(CRC), bladder cancer (BC), kidney cancer (KC), prostate cancer (PC), gastric cancer (GC), and ovarian cancer (OC). Some of these diseases are not clinically correlated (*e.g.*, pre-eclampsia and Parkinson's disease), and therefore, they can serve as a model for evaluating the performance of the artificially intelligent nanoarray in disease diagnosis without disturbances of clinically confounding factors. The other diseases do possess clinical correlation between each other (*e.g.*, lung cancer and pulmonary artery hypertension; colorectal cancer and Crohn's disease), and therefore, they can serve as a model for evaluating performance in disease diagnosis with practical clinical interruptions/effects of clinically confounding factors. Most of the diseases can be categorized into cancerous diseases, inflammatory diseases, neurological diseases, and independent diseases. Therefore, inter- and/or intra-comparison between these groups can evaluate the clinical classification ability of this nanoarray. The mean age of the patient groups was 55 ± 10 years; 423 (52%) of the patients population were male, and 296 (36%) were active smokers. Breath samples were also collected from each of 591 control subjects concurrently enrolled with the patients at each site. The mean age of the control population was 52 ± 8 years; 257 (43%) were males, and 134 (23%) were active smokers. The demographic characteristics of all tested patients and controls in the current study are reviewed in Table 1. All samples were collected according to the Helsinki ethics protocol issued for the study at each of the collaborating institutes and after signed consent of each subject. Detailed information on the clinical design and inclusion and exclusion criteria can be found in the Experimental Methods and Supporting Information (SI), section 1.1, including Tables S1-S12. Detailed information on breath collection and analysis are also in the Experimental Methods.

Breath Analysis with the Artificially Intelligent Nanoarray. During exposure to breath samples, interaction between the VOCs and the organic sensing layer changes the electrical resistance of the sensors. This change recovers to baseline resistance almost immediately at the end of exposure. At this stage, a major caution was taken into account to assure the stability and lack (or minimal) drift of the sensors and/or

Table 1. Demographic Characteristics of Patients and Healthy Controls in the Current Study

group	n	patients		controls	
		age \pm SD ^a	male, n (%)	smoker, n (%)	age \pm SD ^a
lung cancer (LC)	45	67 \pm 09	23 (51%)	44 (98%)	56 \pm 14
colorectal cancer (CRC)	71	66 \pm 10	42 (59%)	09 (11%)	60 \pm 14
head and neck cancer (HNC)	22	62 \pm 12	19 (86%)	13 (59%)	50 \pm 12
ovarian cancer (OC)	48	51 \pm 11	00 (00%)	00 (00%)	47 \pm 09
bladder cancer (BC)	73	69 \pm 11	68 (93%)	53 (68%)	
prostate cancer (PC)	11	66 \pm 08	11(100%)	05 (45%)	66 \pm 12
kidney cancer (KC)	33	65 \pm 13	22 (66%)	15 (45%)	31 (88%)
gastric cancer (GC)	99	63 \pm 12	57 (58%)	26 (27%)	57 \pm 15
Crohn's disease (CD)	41	38 \pm 12	23 (56%)	20 (50%)	55 (34%)
ulcerative colitis (UC)	37	41 \pm 16	20 (56%)	16 (43%)	41 \pm 02
irritable bowel syndrome (IBS)	27	38 \pm 13	08 (32%)	08 (30%)	28 (60%)
idiopathic Parkinson's (IPD)	44	65 \pm 14	23 (53%)	07 (15%)	62 \pm 12
atypical Parkinsonism (PDISM)	16	67 \pm 08	07 (44%)	06 (35%)	19 (51%)
multiple sclerosis (MS)	118	38 \pm 10	42 (36%)	38 (32%)	39 \pm 11
pulmonary hypertension (PAH)	22	48 \pm 12	06 (27%)	12 (54%)	38 \pm 08
pre-eclampsia toxemia (PET)	24	30 \pm 06	00 (00%)	00 (00%)	29 \pm 04
chronic kidney disease (CKD)	82	65 \pm 12	52 (64%)	24 (29%)	46 \pm 02
total	813	55 \pm 10	423 (52%)	296 (36%)	52 \pm 08
				591	257 (43%)
					134 (23%)

^aAge given as mean \pm standard deviation.

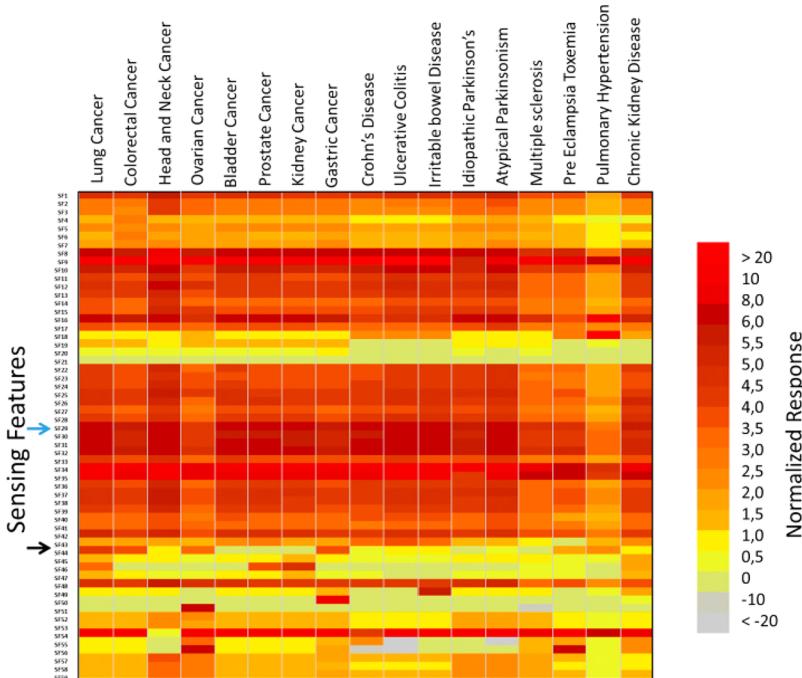


Figure 2. Heat map of 59 stable sensing features, extracted from 20 different nanomaterial-based sensors on the artificially intelligent nanoarray. Each raw datum in the heat map represents the mean responses for each of the 17 diseases tested in this way. Some sensing features (SFs) were more sensitive than others to differences in the breath VOCs. No individual sensing feature was sufficiently informative to discriminate among all the diseases, but the overall response patterns had discriminative potential (columns in the heat map). For details regarding each of the measured sensing features, see SI, Table S13.

sensing features over the entire period of study. The outcome indicated that the sensors were highly stable, with negligible drifts and/or fluctuations being seen during the study (see SI, section 2.1 and Figures S1–S3). This finding negates the possibility of discrimination between different diseases because of drift issues. From each sensing response, four (numerical) sensing features (SFs) were read out: the relative change of sensors' resistance at the peak (beginning), middle, and end of the exposure and the area under the curve of the whole measurement. In total, 59 eligible and stable SFs (SF-01–SF-59) were used for the statistical analysis (Figure 2). For details regarding each of the read sensing features, see SI, Table S13. Figure 2 shows that some sensors were more sensitive to the differences between the VOC patterns of the different disease populations. For example, SF-43 (Figure 2, black arrow, and SI, Figure S4), in which the sensor proved highly discriminative between head and neck cancer and other cancerous disease, whereas discrimination between inflammatory bowel disease and other internal (noncancerous) diseases was not (see SI, Figure S4). Although other sensors were less sensitive, this was indicated by a wide overlap in their responses to the breath samples from different diseases, such as SF-29 (Figure 2, blue arrow). At this stage, a major caution was taken into account prior to statistical analysis. The caution was the exploration of multiple linear regression models for examining and stratifying the effects of possible confounding factors: sex, age, smoking status, and location of sampling site. This analysis showed that, of the 59 studied sensing features, 39 were correlated with age and/or smoking, most of them being negative (*i.e.*, lower signals were obtained from older smokers). Three of the 59 sensing features were correlated with gender. None of the sensing features were correlated with geographical location. Using linear correlations, the raw data were stratified and adjusted,

and a second regression models were applied to ensure the correction was effective. For detailed statistical analysis, refer to the [Experimental Methods](#).

To semiquantify the differences seen in the columns of Figure 2, combinations of sensitive sensors were used to create a series of discriminant factor analysis (DFA) binary classifiers (see [Experimental Methods](#)) to obtain disease breath signatures that allow the different diseases to be distinguished. To ensure valid results free from artifacts or overfitting, we have divided the data set of each analysis as a training and validation set; 77% of each group was selected randomly for the training set, and 23% of each group were omitted as blind samples. The DFA classifier consisted of 120 binary models, each discriminating a pair of the diseases. Thirty randomly chosen samples from each group were used, for this analysis, to ensure uniform sample size. For each binary classifier, 46 samples were used as a training set (23 samples per each of the compared diseases) and 14 randomly chosen independent samples (7 samples per group) were classified in a blind manner (prostate cancer samples were excluded from this specific analysis, due to the small sample size; $n = 11$). The accuracy of the blind analysis of each model was calculated as the total number of samples correctly classified over the total number of independent set samples ($n = 14$). In some cases, the analysis was of low accuracy in discriminating between two groups (*e.g.*, 64% in a comparison of gastric *vs* bladder cancer). In other cases, 100% accuracy was found in 13 different comparisons (*e.g.*, lung *vs* head and neck cancer). The average accuracy of all 120 classifiers was 86%. Figure 3 presents the discriminative power of the nanoarray in terms of accuracy scored in the blind analysis; for the exact sensitivity, specificity, and accuracy of each comparison, please refer to the SI, Table S14. To test whether the discrimination achieved between the different groups was influenced by any

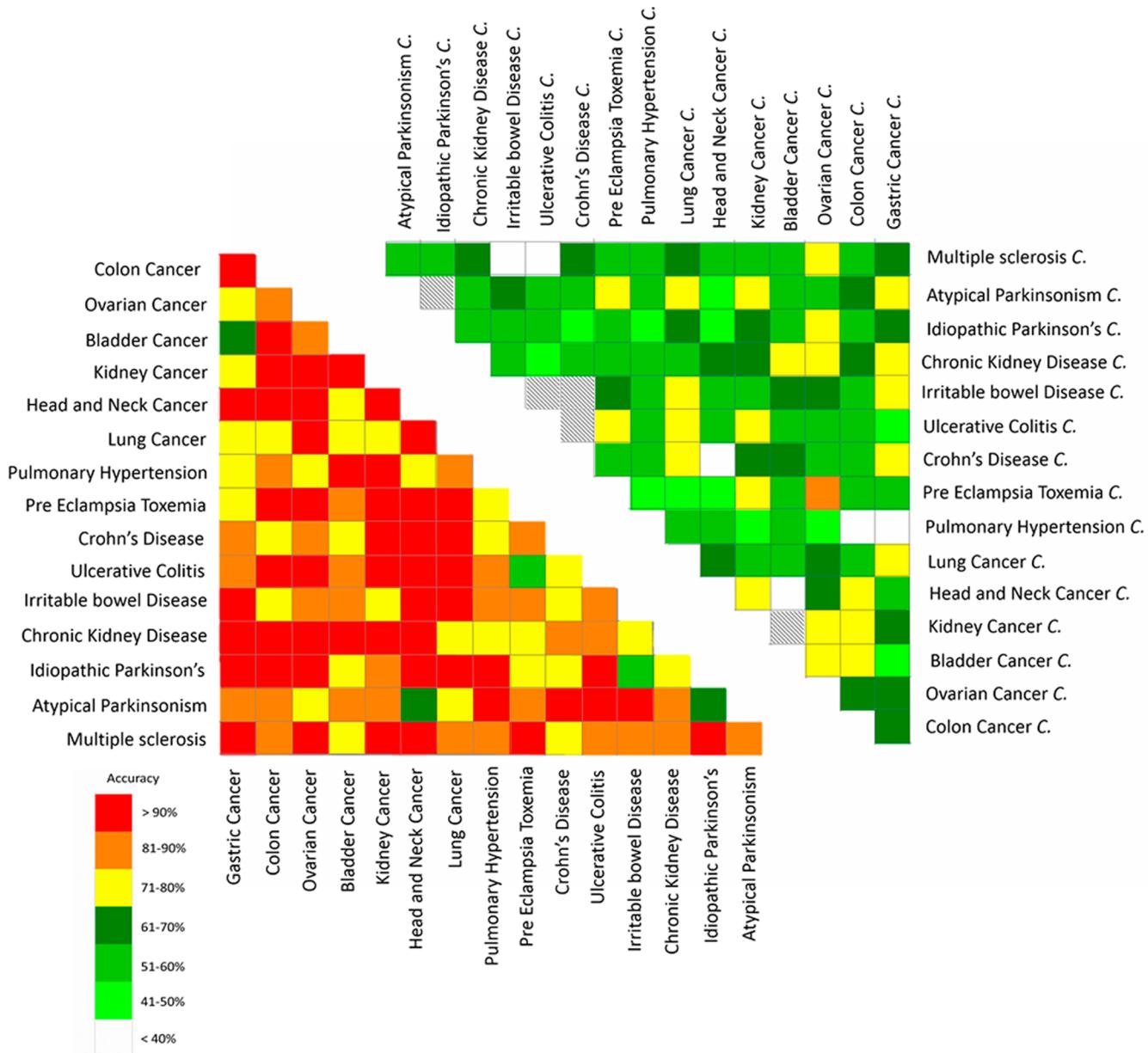


Figure 3. Graphical presentation of the accuracy of the binary DFA classifiers. Each box represents the accuracy achieved in a blind validation of each pair of subject groups. The left heat map gives the results of comparisons between groups of patients, whereas the graph on the right gives the results of the same classifiers applied to the corresponding control groups. The average accuracy was 86% for all disease classifiers (left graph) and 58% for the corresponding control groups (right graph). The letter “C” beside each disease named in the right figure means the “control” group relates to that specific disease.

bias, possibly caused by the confounding factors geography and/or methodology, we applied the exact obtained classifiers that successfully discriminated among the diseases to the corresponding control groups, collected at the same sites under the same conditions and environment. This last analysis resulted in accuracies between 29% for PAH vs CC and 86% for TOX vs OC, yielding an overall average accuracy for this analysis of 58% (Figure 3). (For the exact sensitivity, specificity and accuracy of each comparison, See SI, Table S15). In some cases, two or more diseases shared the same control group, as in (1) Crohn's disease, ulcerative colitis, and irritable bowel syndrome; (2) kidney and bladder cancer; and (3) idiopathic and atypical Parkinsonism. Therefore, the last analysis was not applicable in these cases (Figure 3, hatched boxes). In contrast

to the high accuracy achieved among diseases (86%), the classification of the control samples resulted in random results with a total accuracy of 58%, ruling out the possibility of coincidence. In certain comparisons, the results were higher than the arbitrary classification of the control subjects. Overall, these findings emphasize that the differences in the VOC composition during disease are much more stressed and are more significant than the minor intra-individual differences found among the heterogeneous control groups.

To explore similarities and/or differences among the breath VOCs associated with each disease, hierachal clustering analysis was conducted. In this analysis, the responses of the sensors were clustered and regrouped according to similarities and/or differences in the collective pattern of the VOCs.

Each clustering step represents greater similarities between the profiles, suggesting considerable resemblance among the samples (subjects) of a specific cluster (see **Experimental Methods** for more details). Two important inferences emerged from the results. The first is that the data were not clustered according to possible confounding factors, such as sampling location, racial, and/or ethnic factors (e.g., no resemblance could be seen between multiple sclerosis and Parkinsonian groups; both patients and their corresponding control groups), although both were enrolled and tested in the same department (Carmel Medical Center, Haifa, Israel). In addition, there is no evidence that the samples were clustered due to resemblance in features of sex and/or smoking habits; for example, in the case of pre-eclampsia and ovarian cancer, both groups included only nonsmoking females (see **Figure 4** and **SI**, Figure S5). Second,

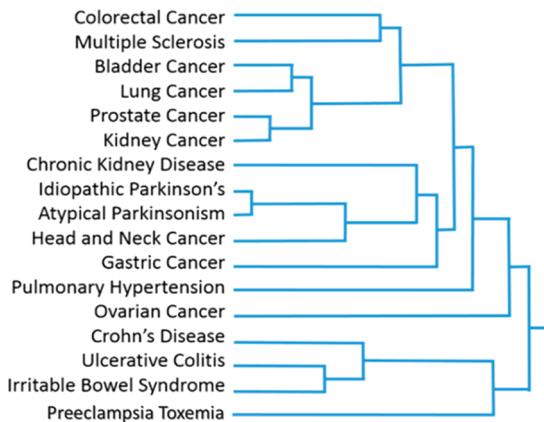


Figure 4. Clustering analysis of the responses of the sensors. Each cluster represents a similar response profile, suggesting considerable resemblance between samples (subjects) in a specific cluster. It is clear that the clustering is not based on any of the potential confounding factors, but there are strong resemblances between subgroups with common pathophysiologies.

there was a strong resemblance between subgroups with common pathophysiologies; for example, a high similarity was found among most of the cancerous disease, as also among diseases associated with increased inflammatory activity (Crohn's disease, ulcerative colitis, and pre-eclampsia), whereas the Parkinsonian-related cases (idiopathic and atypical Parkinsonism) were subgrouped together (**Figure 4**). In a parallel complementary analysis, hierarchical clustering analysis for the corresponding control groups was carried out. The results indicated greater similarities and more homogeneous sensing responses compared with the disease clustering. Figure S5 of the **SI** show that the interclustering distances are shorter than the ones in the disease group clustering. These results support the hypothesis that similarities in pathophysiological processes are expressed in quite similar breath patterns. However, further *in vitro* and *ex vivo* studies are required to support this conclusion.

Complementary Quantitative Chemical Analysis. As mentioned earlier, the artificially intelligent nanoarray analyzes the collective breath VOC patterns in a black-box approach. To identify and quantify the specific VOCs associated with each disease state, a second breath sample obtained from all participants was analyzed by GC-MS. This identified over 150 different VOCs in the different cohorts, but only 35 VOCs were selected for further investigation. The choice was made on

the following criteria: (i) they were common to >70% of the total population (patients and controls); (ii) they were easily identified and verified by the analysis of pure standards; and (iii) they had concentrations in ambient air samples at least 10-fold lower (on average) than in the equivalent breath samples. Owing to the demographic differences between the groups, multiple linear regression for the abundance of each VOCs was first carried out to explore any possible correlation between abundance and the covariates (age, sex, location, and smoking status). The results indicate that the abundances of 15 VOCs were negatively correlated with age and/or smoking; three of them were also correlated with gender. However, there was no significant correlation between the abundance of those VOCs and the sampling site. Therefore, each VOC with significant correlations ($p < 0.05$) was adjusted according to the calculated coefficient corresponding to the confounding element (see **SI**, Table S16). For detailed method of GC-MS analysis, refer to the **Experimental Methods**.

Thirteen different VOCs common to all the examined diseases have been identified in this study and differed significantly ($p < 0.01$) in abundance from the control groups and/or the other diseases. They were 2-ethylhexanol, 3-methylhexane, 5-ethyl-3-methyloctane, acetone, ethanol, ethyl acetate, ethylbenzene, isononane, isoprene, nonanal, styrene, toluene, and undecane (cf. **SI**, Table S17). Although the abundances of these VOCs overlapped among some of the disease states, significant differences were clear between others. For example, nonanal in Crohn's disease, irritable bowel syndrome, and pre-eclampsia was significantly lower than that in other diseases (**Figure 5A**). Undecane in lung, head and neck, and ovarian cancers was markedly higher than that in the remaining groups. The same VOC in Crohn's disease, ulcerative colitis, and pulmonary hypertension was also significantly lower than that in the remaining groups (**Figure 5B**). Isononane, in contrast, among patients with idiopathic Parkinson's disease, atypical Parkinsonism, and pre-eclampsia was significantly higher than that in all other diseases (**Figure 5C**). Similar observations were noted for each of the other 10 VOCs.

Figure 6 shows a heat map of the quantitative GC-MS analysis of patients' breath samples; it is almost impossible to discriminate between the different diseases at the single VOC level (the rows in the heat map), and therefore, a combinatorial approach was taken. The rationale behind this approach is that 13 VOCs could generate thousands of possible combinations in which each combination could be valuable in the discrimination between two or more of the diseases. Indeed, this combinatorial approach shows that the overall combination of 13 VOCs (columns in the heat map) obtained for a specific disease differs from the other diseases (**Figure 6**); that is, in the breath VOC composition of each group examined, the combination of the 13 statistically validated VOCs showed clear differences between the diseases. For example, none of VOC-04, VOC-08, and VOC-11 solely discriminated between the diseases due to the overlap in their individual concentrations. However, the combination of these three VOCs showed clear distinction between colorectal cancer and ulcerative colitis, as well as between head and neck cancer and multiple sclerosis. Adding VOC-02 and VOC-12 to these three VOCs, the discriminative power increases, making it possible to discriminate between additional diseases, including lung cancer, pulmonary artery hypertension, and Parkinson's disease. These results clearly indicate that assessment of wide subsets (patterns) of the 13 VOCs

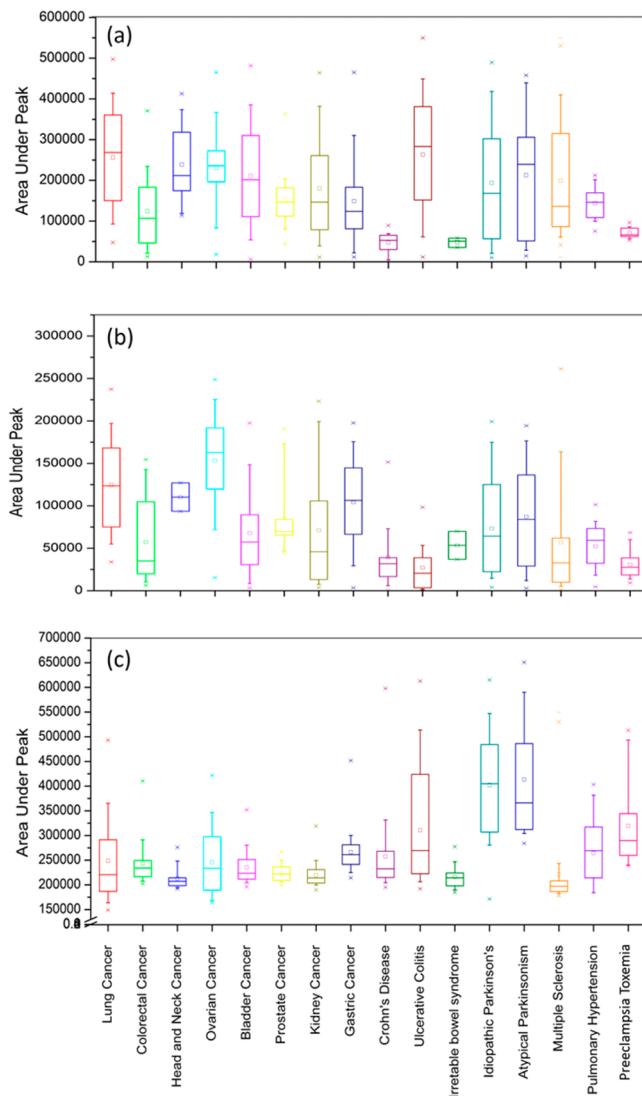


Figure 5. GC-MS analysis of the breath samples. The area under the peak (abundance) measured in the different diseases of three representative VOCs: (a) nonanal, (b) undecane, and (c) isononane. The whisker boxes present first quartile, third quartile, median (line), and average (square); the bars represent the 10% and 90% points, whereas the dots represent minimal and maximal readings.

gives better discrimination between the diseases we have explored.

For the sake of comparison, the discriminatory power of the 13 VOCs in Figure 6 was examined in the context of the different control groups, which showed that the number of cases of significant differences ($p < 0.01$) in VOC abundance among the (control) groups being assessed was 50% lower than in the case-related disease groups. In fact, there were significant differences ($p < 0.01$) in 177 of the 858 (21%) of the binary comparisons among the control groups collected concomitantly with the disease groups. On the other hand, there were significant differences ($p < 0.01$) in 760 of the 1768 (43%) of the binary comparisons between the disease groups. This demonstrates the intra-individual differences in the breath VOC composition of the controls, but much more obvious and remarkable ones were found among patients with different diseases. Nevertheless, and in spite of the intra-individual VOC differences in the control groups, it is clear that each disease has

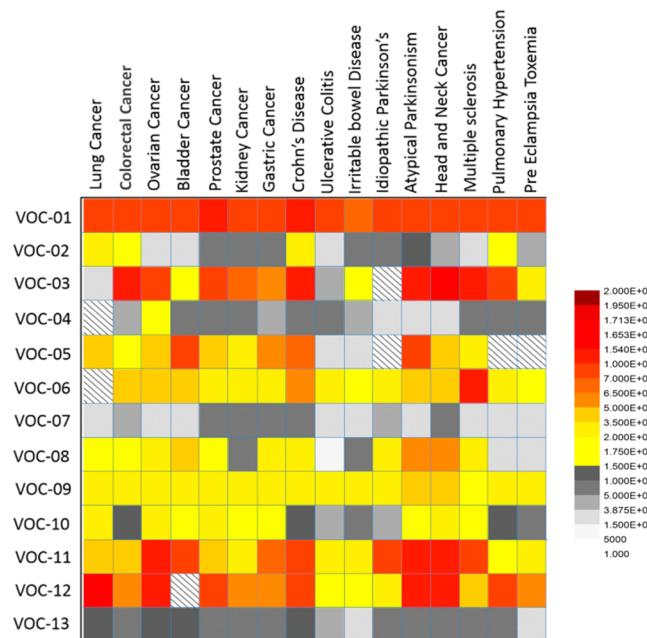


Figure 6. Heat map of the GC-MS analysis of patients' breath samples. The average of each of the 13 VOCs is given on the color scale. Hatched boxes are cases in which the VOC was found in <70% of the samples of a specific group of patients. The VOCs are VOC-01, 2-ethylhexanol; VOC-02, 3-methylhexane; VOC-03, 5-ethyl-3-methyloctane; VOC-04, acetone; VOC-05, ethanol; VOC-06, ethyl acetate; VOC-07, ethylbenzene; VOC-08, isononane; VOC-09, isoprene; VOC-10, nonanal; VOC-11, styrene; VOC-12, toluene; and VOC-13, undecane. At the individual VOC level, it becomes almost impossible to distinguish between different diseases (rows in the heat map). However, the overall VOC composition (columns in the color map) is quite distinctive.

a remarkably distinctive VOC composition of the breath that distinguishes it from both the controls and other disease states.

DISCUSSION

Analysis with the Artificially Intelligent Nanoarrays.

The binary classifiers of the breath samples collected from the various disease states achieved very high accuracy in most cases, supported by two independent validation procedures, viz. the leave-one-out cross-validation method (applied to the training set) and blind analysis of 14 samples for each classifier. A minimal number (2–4) of sensors was also used for each comparison, reducing the chance of overfitting the results. In contrast to the high accuracy achieved among diseases (86%), the classification of the control samples ruled out the possibility of coincidence and/or external biases. Of special importance, results from the artificially intelligent nanoarrays support the hypothesis that similarities in pathophysiological processes are expressed in quite similar breath patterns. The results also indicated that the adjustment for confounding factors was successful. The subgroups were not clustered according to similarities in demographic features or geographical location, which also stresses that the artificially intelligent nanoarray analysis is less sensitive to possible confounding factors since we have seen in some cases trends in the control groups that were like those seen among the diseases.

These observations could be explained by the fact that the artificially intelligent nanoarray does not rely on selective or a lock-and-key approach in the detection of specific VOCs in the

biomarkers. Indeed, the nanoarray relies on sensors that have different chemistries. In this case, all exhaled VOCs are adsorbed onto the surface of each sensor. Due to the different chemistry of each sensor, however, a combination of the 13 VOCs could have higher affinity to the sensor's surface than others. In this way, different combination of VOCs (of the 13 VOCs) are detected by different sensors. With the help of artificial intelligence methods, combination of these differences can be expressed and attributed to specific diseases.^{6,10,37}

To date, some of the classifiers gave relatively poor accuracy, which will be ineffective in future clinical use. Nevertheless, these results may well be improved by using tailor-made sensors that are more sensitive to specific VOC profiles.^{6,10} New-generation sensors with modified and more diverse organic coatings may be fabricated to suit particular VOC composition(s). Their tailor-made characteristics could be determined by repeated exposure and training phases using artificial gas samples consisting of the targeted VOC mixture.^{36,59} For full and efficient operation in real clinical settings, the artificially intelligent nanoarray needs to be increasingly trained using known clinical samples to build up a consistent and reliable database of reference. It can then recognize new samples by comparing disease-related VOC patterns to those already in its database. On this issue and other limitations involved in the current study, refer to SI, section 2.4.

Analysis with the GC-MS. Hundreds of different exhaled VOCs have been reported in the literature.^{2–5,8,9,20,60} In our study, 150 different VOCs were identified in their different groups, in agreement with previously published studies (ranging from 100 to 300 VOCs per subject).^{8,57} Thirteen of these VOCs could discriminate between different diseases. These VOCs are well-known, documented in the literature, and each single VOC has been reported in different disease states. For example, nonanal was linked to ovarian cancer,²⁵ inflammatory bowel disease,⁴⁵ breast cancer,⁶¹ esophageal and gastric adenocarcinoma.⁶² As another example, isoprene could be linked to chronic liver disease,⁶³ kidney disease,⁵¹ diabetes,⁶⁴ and others. These results support our finding that no single VOC can discriminate between different diseases. The use of VOC patterns in exhaled breath samples becomes a more realistic option for discriminating between different disease states.

Regression models applied to the raw GC-MS data showed that the abundance of exhaled VOCs was affected by some common confounding factors. A number of the VOCs was affected by age and/or smoking habits (e.g., 2-ethylhexanol, 3-methylhexane, 5-ethyl-3-methyloctane, acetone, ethanol, ethyl acetate, ethylbenzene, isononane, isoprene, nonanal, styrene, toluene, and undecane), whereas three of them were also affected by the gender of the subject (isononane, nonanal, and undecane). This effect stemming from the first part of the VOCs could be explained by the relationship between the anatomical and physiological changes in the respiratory system and circulation associated with aging and/or smoking injury.⁶⁵ It includes stiffness and degeneration of the elastic fibers, fibrosis, aging-associated destruction of lung parenchyma, emphysema, and chronic bronchitis, mainly among smokers.⁶⁶ These alterations could easily affect the diffusion of VOCs through the blood–air barrier by altering the thickness or permeability of the epithelium (the so-called membrane conductance) or by reducing the total surface area of the membrane.⁶⁶ These factors could easily alter the flux, according to Fick's first law, affecting the diffusion of gases in the exhaled air, eventually reducing/stressing the expression and/or

concentrations of a wide range of the exhaled VOC components.⁵ The effect stemming from the second part of the VOCs might be attributed to hormonal or structural gender-related differences.⁶⁷ However, it remains unclear why in the case of 3-methylhexane, and despite the strict adjustments that were made, an unexplained trend was found among the control groups. The concentration of 3-methylhexane was lower in the multiple sclerosis group than any other disease group and lower in the corresponding control group than any of the other control group. It is believed that these rare trends are due to a range of factors, such as small differences between the different clinical teams in using the sampling protocol and among the technical properties of the analytical instruments during analysis. However, it is noteworthy that the multiple sclerosis patients and their corresponding control group had been enrolled and sampled in the same location as the Parkinsonian patients, with the same equipment and by the same staff. Therefore, it is concluded that the measured concentrations should be corrected in future applications of breath testing. These results are contrary to many previous reports in the breath analysis field, which commonly reported that only a small minority of breath VOCs can be correlated with common confounding factors.^{68–70} The main reasons for this could include the following: (i) in studies consisting of matched groups, these features had been neglected and left untested; and/or (ii) owing to the small size of study groups, statistically significant correlations would not have been obtained. These conclusions will be considered as empirical findings—as the principal origins and specific pathways of our VOCs—that remain poorly understood, although intensive experimental work and theoretical considerations have been discussed elsewhere.^{3,5,8} Further *in vitro* and *ex vivo* studies, in which, for example, healthy cells/tissues are stimulated with different pathophysiological stimuli to monitor changes in their VOC patterns will be required to understand and further strengthen our conclusions. For further details on the limitations involved in the current proof-of-concept study, refer to the SI, section 2.4.

Summary and Conclusions. The present study reports on an artificially intelligent nanoarray based on molecularly modified gold nanoparticles and random network of single-wall carbon nanotubes for noninvasive diagnosis and classification of 17 different diseases from exhaled breath. The nanoarray was used for the practical evaluation of 1404 subjects in nine clinical settings worldwide. Blind experiments with the artificially intelligent nanoarray showed that 86% accuracy could be achieved, allowing discrimination between each pair of the diseases, and that each disease has its own unique volatile molecular print compared to both healthy controls and other diseases. Consequently, the presence of one disease would not screen out others—a prerequisite for developing a new generation(s) of biomedical devices for personalized diagnosis and classification of diseases in a noninvasive, inexpensive, and portable manner. The artificially intelligent nanoarray had a low or no vulnerability to clinical and demographical confounding factors. The findings by nanoarray were examined by an independent analytical technique, GC-MS. This analysis found 13 exhaled VOCs associated with various diseases, and their composition differs from one disease to another, thereby validating the nanoarray results. While further and larger translational studies are required to validate these findings, this work provides a shuttling pad for *in statu nascendi* “volatolomics” field (the omics of volatile biomarkers), as well as a method for obtaining affordable, easy-to-use, inexpensive, and miniaturized tools for

personalized screening, diagnosis, and follow-up of a range of diseases.

EXPERIMENTAL METHODS

Fabrication of the Nanoarray. Breath samples obtained by each subject were analyzed by an array of cross-reactive sensors that relates to two main chemiresistive categories: (i) monolayer-capped gold nanoparticles (core diameter: 3–4 nm), and (ii) an organically functionalized random network of single-walled carbon nanotubes (SWCNTs). The gold nanoparticles were synthesized as previously described^{32,71,72} and drop-casted onto semicircular microelectronic transducers until a resistance of several $M\Omega$ had been reached. The microelectronic transducers consisted of 10 pairs of circular interdigitated (ID) gold electrodes on silicon with 300 nm thermal oxide (Silicon Quest International, Nevada, USA). The outer diameter of the circular electrode area was 3 mm, and the gap between two adjacent electrodes and the width of each electrode were both 20 Ω m; the sensors were molecularly modified with different sensing layers, such as dodecanethiol, 4-chlorobenzene methanethiol, and several others. The SWCNT sensor was based on an electrically continuous random network of SWCNTs formed by drop-casting a solution of SWCNTs (ARRY International LTD, Germany; ~30% metallic, ~70% semiconducting, average diameter 1.5 nm, 7 mm) in dimethylformamide (DMF, from Sigma-Aldrich Ltd., >98% purity) onto the prepared electrical transducers. After deposition, the device was slowly dried overnight under ambient conditions to enhance the self-assembly of the SWCNTs and evaporate the solvent. This procedure was repeated until a resistance of 100 $K\Omega$ to 10 $M\Omega$ was reached. The microelectronic transducer for the SWCNT sensor consisted of 10 pairs of 4.5 mm wide interdigitated Ti/Pd electrodes on silicon with 2 μ thermal oxide (Silicon Quest International, Nevada, USA). The SWCNT sensor was organically functionalized with PAH derivatives. For details regarding each of the sensors, see SI, Table S13. After fabrication, each sensor went through a characterization procedure in which it was exposed to several different concentrations of compounds usually found in breath (e.g., isopropyl alcohol, 2-ethyl hexanol, water vapor, and others), in a range of 10 or more parts per billion (ppb) – several part per millions (ppm). Features including, but not only, sensitivity, selectivity, repeatability, and stability of each sensor were carefully studied, then the information from these experiments offered the choice of the most sensitive, stable, and repeatable sensors to be used in the study.

Design of the Clinical Study. Breath samples collected between January 2011 and June 2014 were subjected to meta-analysis; they were provided by collaborators of our research group (LNBD; Technion-IIT) with 14 departments in nine clinical centers worldwide. Originally, breath samples had been collected at each site for different scientific purposes. Nevertheless, the same collection equipment and the strict procedure were used by all participating centers to ensure the samples were comparable. However, it should be noted that different sampling locations had different inclusion/exclusion criteria, such that in some cases the different subpopulations were not matched in terms of their demographic characteristics. Analysis was made on two breath samples from each of the 813 patients diagnosed with one of the following diseases: chronic kidney failure (CKD) ($n = 82$) (Poria Hospital, Tiberias, Israel); idiopathic Parkinson's disease (IPD) and atypical Parkinsonism (PDISM) ($n = 60$) (Carmel Medical Center, Haifa, Israel); multiple sclerosis (MS) ($n = 118$) (Carmel Medical Center, Haifa, Israel); inflammatory bowel disease (IBD), including Crohn's disease (CD) ($n = 41$), ulcerative colitis (UC) ($n = 37$) and irritable bowel syndrome (IBS) ($n = 27$) (Rambam Medical Center, Haifa, Israel); pulmonary arterial hypertension (PAH) ($n = 22$) (French National Referral Center for PAH, Antoine-Béclère Hospital, Paris, France); pre-eclampsia in pregnant women (PET) ($n = 24$) (Nazareth English Hospital, Nazareth, Israel, and Emek Medical Center, Afula); head and neck cancer (HNC) ($n = 22$) (Carmel Medical Center, Haifa, Israel); lung cancer (LC) ($n = 45$) (Baptist Cancer Institute (BCI), Jacksonville, FL, USA); colorectal cancer (CRC) ($n = 71$) (Faculty of Medicine, Riga East University Hospital, Riga, Latvia); bladder cancer

(BC) ($n = 73$), kidney cancer (KC) ($n = 33$) and prostate cancer (PC) ($n = 11$) (Bnai-Zion Medical Center, Haifa, Israel); gastric cancer (GC) ($n = 99$) (Faculty of Medicine, Riga East University Hospital, Riga, Latvia); and ovarian cancer (OC) ($n = 48$) (Department of Oncology, First Affiliated Hospital of Anhui Medical University, Hefei, China). The mean age of the patient groups was 55 ± 10 years, 423 (52%) being male, which included 296 (36%) active smokers. Breath samples were also collected from 591 control subjects enrolled concurrently with the patients at each site. The control population mean age was 52 ± 8 years; 257 (43%) were males, of which 134 (23%) were active smokers. Demographic characteristics of all these subpopulations are summarized in Table 1. Samples were collected according to the Helsinki ethics protocol issued for study at each of the collaborating institutes and after the signed consent of each subject. Over 95% of the patients had been enrolled before therapeutic intervention (i.e., *de novo*); for ethical reasons, the remaining cases were sampled after treatment had begun. Detailed information on the clinical design, inclusion, and exclusion criteria can be found in the SI, section 1.1.

Collection of the Breath Samples. Briefly, the breath collection procedure started with a 3–5 min “lung wash” during which the participant inhaled through both a charcoal (ABEK) and a bacterial filter to remove 99.99% of exogenous VOCs in the ambient air. The subject then exhaled through a device that had two separate ports; one port directed the first 150 mL of exhaled breath, mainly dead space air, to a waste plastic bag, and the rest (the end tidal volume) was directed into a 750 mL chemically inert Mylar sampling bag. The process was repeated within a few minutes to obtain two samples from each subject. The content of each Mylar bag was transferred using a simple vacuum air pump into a Tenax TA and Carboxen-1018 glass adsorbent tube (Sigma-Aldrich Ltd.) or two-bed ORBO 420 Tenax TA sorption tubes (Sigma-Aldrich, St. Louis, MO, USA). For quality control, air samples were collected by pumping ambient air from the collection room through a sorbent tube for 7 min at 100 mL/min. The tubes were sealed and stored in refrigerators at 4 °C pending analysis. One sample was used for chemical analysis using GC-MS and the other for nanoarray pattern recognition analysis. As a precaution, only subjects who had *not* consumed food, drink, or tobacco at least 2 h before sampling were included.

Breath Analysis with the Artificially Intelligent Nanoarray. To transfer VOCs trapped by the absorption materials, the samples underwent thermal desorption (at 250 °C) in an autosampler thermal desorption system (TD20; Shimadzu Corporation, Japan) before they were temporarily stored in a stainless steel VICI sample loop (Valco Instruments Co. Inc.) at 150 °C. In parallel, the chamber containing sensors was kept under vacuum conditions (~30 mTorr), until the sample was directed into the chamber; the remaining volume was filled with N₂ (99.999%) at atmospheric pressure. A Keithley data logger device (model 2701 DMM) was used to sequentially acquire resistance readings from the nanoarray for 5 min in vacuum prior to exposure (baseline), followed by 5 min of breath sample that filled the chamber, which in turn was followed by another 5 min of sensors recovery starting with a vacuum chamber. The whole system was controlled by a custom-built LabView program.

Calibration and Signal Analysis. During exposure of the sensors to breath samples or calibration gas, interaction between the VOCs and the organic sensing layer results in a change in resistance. Interactions between the SWCNTs/GNPs and VOCs are physical, mostly relying on van der Waals interactions.^{3,37} Baseline resistance recovers almost immediately with the evacuation of the sample from the chamber. From each response, four (numerical) sensing features were read out: the relative change of sensor's resistance at the peak (beginning), middle, and end of the exposure, as well as the area under the curve of the whole measurement. For example, the middle sensing feature, of a given exposure, was calculated by the following equation:

$$\text{SF} = (\text{sensor resistance at the middle of the exposure} - \text{baseline resistance}) / \text{baseline resistance}$$

To supervise the sensor's operational efficiency during the experiment and to overcome drift in sensors' response, a fixed calibration gas mixture, containing 11.5 ppm isopropyl alcohol, 2.8 ppm trimethylbenzene, and 0.6 ppm 2-ethylhexanol, was exposed to the sensors on a daily basis. The raw signals of the breath samples' response were normalized by their parallel response to the calibration gas on the same day. We also used a humidity compensation method to reduce the effect of water vapor present at higher levels in samples collected on Tenax-Carboxen trap tubes, due to its ability to adsorb water vapor.⁵⁹ Features were extracted only from 20 sensors that were functional and stable throughout the entire analysis. Characteristics and sensing features of the eligible sensor are detailed in **Table S13**.

Multiple linear regression models were applied to the extracted sensing features to test and correct the associations between covariates (age, sex, location, and smoking status) and the numerical outcome. To choose the most sensitive sensors relevant to a given comparison and obtain the most selective models for the targeted VOCs, statistical analysis was used to indicate the *F* ratio of each sensor (*i.e.*, its power of discrimination). DFA was then applied to the data (highest ranking SF) to test the ability of the sensors to distinguish the different diseases from one another. DFA is a supervised *linear* method supplied with the classification information for every measurement in the training set. Using a linear combination of the input variables E_n , DFA finds new orthogonal axes (canonical values) to minimize the variance within each given class and maximize variance between two classes. To prevent overfitting of the data, two to four sensing features were used to build a discriminatory model, stressing a ratio of 1:10 of features to samples. Thirty samples (or less in cases with a limited number of patients) were randomly chosen and used to obtain binary classifiers using DFA. From each group, seven samples were first excluded for blind validation of the acquired classifier; the remaining samples were used as the training set. Each of the samples were then classified accordingly and we determined whether the result turned out to be true positive (TP), false positive (FP), true negative (TN), or false negative (FN). The classification accuracy was identified by calculating the sensitivity (TP/TP + FN), specificity (TN/TN + FP), and accuracy (TP + TN/sample size) of the constructed model.

Both cross-validation methods and blind tests were applied to validate the results. The responses and selected sensing features of 20 eligible sensors were then normalized according a daily calibration system. Control experiments to test the reproducibility of the sensing responses and behavior were carried out over periods ranging from 6 months to 3 years. Some day-to-day fluctuations were evident in the calibration experiments, but their magnitude (noise) was negligible. For detailed information regarding signal processing, please refer to refs 23–25 and 53.

Chemical Analysis Using GC-MS. In general, GC-MS analysis is divided into two steps. In this first step, the sample goes through the GC capillary column, in which the velocity of a molecule is correlated to its chemical properties; that is, different molecules will exit the column at different times (their retention times). Separation of the mixture allows the MS to capture, ionize, accelerate, deflect, and detect the ionized molecules separately. The MS does this by breaking each molecule into ionized fragments, which detects using their mass to charge ratio and compares it to a given library. The final output contains a chromatogram, which relates the retention time to the abundance of each compound. The area under the curve for a specific compound is proportional to its concentration in the mixture. One of the two breath samples from each subject was analyzed using GC-MS to determine the chemical nature and composition of his/her exhaled breath. For this analysis, GC-MS (GCMS-QP2010; Shimadzu Corporation, Japan), combined with a thermal desorption system (TD20; Shimadzu Corporation, Japan), was used. The following oven temperature profile was set: (a) 10 min at 35 °C; (b) a 4 °C/min ramp until 150 °C; (c) a 10 °C/min ramp until 300 °C; and (d) 5 min at 300 °C. An SLB-5 ms capillary column (Sigma-Aldrich Ltd.) with 5% phenyl methyl siloxane (30 m length, 0.25 mm internal diameter, and 0.5 μm thickness) was used. A splitless injection mode was used for 2 min at 30 cm/s constant linear speed and 0.70 mL/min column flow. Prior to analysis, the tubes were conjugated to a 30 mL/min flow of

pure helium to reduce their humidity. GC-MS chromatogram analysis was undertaken with the GC-MS solutions version 2.53SU1 Postrunanalysis program (Shimadzu Corporation). Student *t*-test and/or nonparametric Wilcoxon tests are used to determine statistical differences of the VOCs' abundance between the research groups. Due to major technical difficulties and repair of the instrument, the breath samples collected from CKD patients were analyzed in a different manner and with different technical parameters from the rest of the study cohort and, therefore, were excluded from the GC-MS analysis.

Statistical Analysis for the Artificially Intelligent Nanoarray.

Student *t*-test was used to determine statistical differences between the clinical/demographical features between study groups (including age, sex, smoking habits, *etc.*); $p < 0.05$ was used as a cutoff to determine significant differences. Multiple linear regression models were used to explore and stratify the effects of possible confounding factors, including sex, age, smoking status, and location of sampling site. Sensors' responses that could be correlated with any of the tested variables ($p < 0.05$) were adjusted according to the relevant coefficient factors. After data correction, a second regression model was run to ensure the correction was effective. Statistical differences in the responses of the sensors to different study groups were tested using nonparametric Wilcoxon test due to the non-normal distribution; $p < 0.01$ was used to determine statistical validation. DFA was carried out detailed in **SI**, section 2.4. The validation of the each obtained model was tested by blind analysis of 14 samples (seven samples from each of the compared groups). Each of the samples was then classified accordingly to determine whether the result was a true positive, false positive, true negative, or false negative. Classification accuracy was checked by calculating the sensitivity (TP/TP + FN), specificity (TN/TN + FP), and accuracy (TP + TN/sample size) of the constructed model. Hierarchical clustering analysis was run on individual sensors' responses in all the tested populations. Each clustering step indicates greater similarities between the profiles, suggesting considerable resemblance among the samples (subjects) of a specific cluster.

Statistical Analysis for the GC-MS. After identification and quantification of the exact composition of exhaled breath samples, multiple linear regression models were used to test the effects of confounding factors (sex, age, smoking status, and location of sampling site) on the concentrations of VOCs. Once a significant correlation had been noted with any of the tested variables ($p < 0.05$), we applied a data adjustment method according to the relevant coefficient factors. After data correction, a second regression model was obtained to ensure the correction was effective. The Shapiro-Wilk test was used to check whether the concentrations of the reported VOCs were normally distributed. Thus, Student *t*-test and/or nonparametric Wilcoxon tests could then be used to determine statistical differences of VOCs' abundance between the research groups; $p < 0.01$ was used as a cutoff for statistical validation.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acsnano.6b04930](https://doi.org/10.1021/acsnano.6b04930).

Detailed methods section, additional results and figures, and an appendix on the inclusion and exclusion criteria of the study population ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

*E-mail: hhossam@technion.ac.il.

ORCID

Douglas Johnson: [0000-0001-6584-1327](https://orcid.org/0000-0001-6584-1327)

Hossam Haick: [0000-0002-2370-4073](https://orcid.org/0000-0002-2370-4073)

Author Contributions

R.J., M.A., Y.Y.B., A.G., H.I., S.K., S.B., A.B.H., L.H.S., L.G.M., I.L., A.M., S.B., J.F., S.C.K., F.N., M.H., Z.B., A.K., R.W., F.P.,

- (67) Jain, R. B. Distributions of Selected Urinary Metabolites of Volatile Organic Compounds by Age, Gender, Race/Ethnicity, and Smoking Status in a Representative Sample of U.S. Adults. *Environ. Toxicol. Pharmacol.* **2015**, *40*, 471–479.
- (68) Lechner, M.; Moser, B.; Niederseer, D.; Karlseder, A.; Holzknecht, B.; Fuchs, M.; Colvin, S.; Tilg, H.; Rieder, J. Gender and Age Specific Differences in Exhaled Isoprene Levels. *Respir. Physiol. Neurobiol.* **2006**, *154*, 478–483.
- (69) Phillips, M.; Cataneo, R. N.; Greenberg, J.; Gunawardena, R.; Naidu, A.; Rahbari-Oskoui, F. Effect of Age on the Breath Methylated Alkane Contour - A Display of Apparent New Markers of Oxidative Stress. *J. Lab. Clin. Med.* **2000**, *136*, 243–249.
- (70) Schwarz, K.; Filipiak, W.; Amann, A. Determining Concentration Patterns of Volatile Compounds in Exhaled Breath by PTR-MS. *J. Breath Res.* **2009**, *3*, 027002.
- (71) Dovgolevsky, E.; Konvalina, G.; Tisch, U.; Haick, H. Monolayer-Capped Cubic Platinum Nanoparticles for Sensing Nonpolar Analytes in Highly Humid Atmospheres. *J. Phys. Chem. C* **2010**, *114*, 14042–14049.
- (72) Dovgolevsky, E.; Tisch, U.; Haick, H. Chemically Sensitive Resistors Based on Monolayer-Capped Cubic Nanoparticles: Towards Configurable Nanoporous Sensors. *Small* **2009**, *5*, 1158–1161.