Analysis of exhaled breath for diagnosing head and neck squamous cell carcinoma: a feasibility study

M Gruber[1,3], U Tisch[2,3], R Jeries[2], H Amal[2], M Hakim[2], O Ronen[1], T Marshak[1], D Zimmerman[2], O Israel[2], E Amiga[2], I Doweck*,[1] and H Haick*,[2]

[1]The Department of Otolaryngology Head and Neck Surgery, Carmel Medical Center, Haifa 34362, Israel and [2]The Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion – Israel Institute of Technology, Haifa 3200003, Israel

Background: Squamous cell carcinoma of the head and neck (HNSCC) are wide-spread cancers that often lead to disfigurement and loss of important functions such as speech and ingestion. To date, HNSCC has no adequate method for early detection and screening.

Methods: Exhaled breath samples were collected from 87 volunteers; 62 well-defined breath samples from 22 HNSCC patients (larynx and pharynx), 21 patients with benign tumours (larynx and pharynx) and 19 healthy controls were analysed in a dual approach: (i) chemical analysis using gas chromatography/mass spectrometry (GC–MS) and (ii) breath-print analysis using an array of nanomaterial-based sensors, combined with a statistical algorithm.

Results: Gas chromatography/mass spectrometry identified ethanol, 2-propenenitrile and undecane as potential markers for HNSCC and/or benign tumours of the head and neck. The sensor-array-based breath-prints could clearly distinguish HNSCC both from benign tumours and from healthy states. Within the HNSCC group, patients could be classified according to tumour site and stage.

Conclusions: We have demonstrated the feasibility of a breath test for a specific, clinically interesting application: distinguishing HNSCC from tumour-free or benign tumour states, as well as for staging and locating HNSCC. The sensor array used here could form the basis for the development of an urgently needed non-invasive, cost-effective, fast and reliable point-of-care diagnostic/screening tool for HNSCC.

Head and neck cancer (HNC) refers to a group of diverse tumours in the region of the head and neck (Pai and Westra, 2009). Most HNCs (>90%) are squamous cell carcinoma of the head and neck (HNSCC) that arise from the mucosa lining of the upper aero-digestive tract (Pai and Westra, 2009; Marur *et al*, 2010). Prognosis and treatment depend very strongly on stage and site of the disease. The overall 5-year survival rate is high (90%) for stage I disease, 70% for stage II disease and 30–50% for locally advanced disease (Argiris *et al*, 2008; Haddad and Shin, 2008). The different subsites of the primary tumour also affect the overall survival. Early, differential diagnosis including identification of disease stage and site is therefore highly important. Conventional diagnosis and staging of

HNSCC is based on identification of macroscopic disease by physical examination, radiological studies and tissue sampling. However, to date, less than half of the patients are diagnosed with early disease, due to the shortcomings of the currently available diagnostic and screening methods (Argiris *et al*, 2008).

Non-invasive profiling of exhaled volatile organic compounds (VOCs; i.e., organic compounds that have a high vapour pressure) could provide a fast and convenient alternative solution for early HNSCC diagnosis and screening. Over the past two decades, the analysis of VOCs has attracted much research interest due to its potential for rapidly diagnosing a variety of diseases, for example, cancer, kidney disease, multiple sclerosis or Parkinson's disease,

using either methods of chemical analysis and compound identification, or breath-print analysis based on the collective output of sensor arrays (Gordon *et al*, 1985; O'Neill *et al*, 1988; Mendis *et al*, 1994; Phillips *et al*, 1994; Phillips *et al*, 1999; Miekisch *et al*, 2004; Amann *et al*, 2007; Barash *et al*, 2009; Peng *et al*, 2009; Peng *et al*, 2010; Tisch and Haick, 2010a; Shuster *et al*, 2010; Song *et al*, 2010; Hakim *et al*, 2011; Ionescu *et al*, 2011; Tisch *et al*, 2012; Broza and Haick, 2013; Konvalina and Haick, 2014).

Several pilot studies with small patient cohorts have addressed the possibility of developing a VOC-based breath test for HNCs. Schmutzhard *et al* (2008) were the first to show that exhaled VOCs may be used to distinguish HNSCC patients both from high-risk (i.e., heavy smokers and drinkers) and low-risk healthy controls. They determined the concentration profiles of several HNC-specific VOCs, mainly hydrocarbons (alkanes, alkenes, alcohols, ketones and organic acids) by proton-transfer reaction mass spectrometry. However, the compounds were not identified by name. García *et al* have recently identified seven VOCs, among them ethanol, as possible markers of laryngeal carcinoma, using solid-phase microextraction (SPME) combined with gas chromatography/mass spectrometry (GC–MS). However, they did not attempt quantification, and their study group comprised only 11 laryngeal carcinoma patients and 10 healthy controls (García *et al*, 2014). Leunis *et al* (2013) have chosen the alternative approach and have attempted identifying the breath-print of HNSCC from exhaled breath samples, using a commercial sensor array and statistical pattern analysis software. They could distinguish HNSCC patients from a control group of patients having unspecified benign conditions with 90% sensitivity and 80% specificity. Hakim *et al* (2011) have recently demonstrated the feasibility of an overall HNC breath test, using a dual approach that included a rigorous chemical analysis based on SPME with GC–MS, as well as a breath-print analysis using a custom-made array of nanomaterial-based sensors. In this way a biologically diverse group of HNC patients (different histology and tumour sites) could be distinguished from a group of healthy controls and from a second control group of lung cancer patients.

In this study, we will demonstrate the feasibility of exhaled breath analysis for a more specific and clinically interesting application: identifying, locating and staging HNSCC of the larynx and pharynx. For this purpose we analysed the breath samples of a mixed population of tumour-free subjects and subjects with benign or malignant tumours of the larynx and pharynx, using a dual approach combining chemical analysis (GC-MS coupled with an improved sample pre-concentration method that is suitable for high-throughput sampling) and sensor-based breath-print analysis (with an improved sensor array).

PATIENTS AND METHODS

Patients. Breath samples were collected at the Otolaryngology Head and Neck Department, Carmel Medical Center, Haifa, Israel, from 87 female and male volunteers after obtaining written informed consent. These included 68 patients with benign or malignant head and neck lesions and 19 healthy controls. The healthy controls were recruited among the patients' accompanying persons, usually their spouses, in order to match them to the patients with regard to age and lifestyle. In this way, however, the control and patient groups could not be gender-matched, because benign and malignant lesions of the head and neck occur predominantly in men. The healthy volunteers were not aware of any disease and did not undergo medical examination. The following exclusion criteria were applied to all 87 volunteers before sample collection: medical history of any malignancy as well as any former oncological treatment, age <18 years, an active infectious

disease, present antibiotic treatment, pregnancy or lactation. The 68 patients underwent pertinent anamnesis, physical examination and radiological studies when indicated, in view of the different lesions. Flexible nasolaryngoscopy was a mandatory step in the preliminary assessment. Biopsies were then taken from all 68 participating patients for tissue diagnosis. Breath samples were collected before taking biopsies; enrolment in the study did not delay the biopsy or interfere with the management protocol in any case. Oropharyngeal malignant lesions were assessed for human papillomavirus (HPV) status by immunohistochemistry and all samples were HPV negative.

Results of the tissue diagnosis: malignant disease (27), severe dysplasia/carcinoma *in situ* (4), mild to moderate dysplasia (2) and benign lesions (31). Four patients had not been diagnosed at the time of analysis of the breath samples. After the breath collection and diagnosis, and prior to the analysis we have excluded 25 patients with ill-defined or heterogeneous symptoms, that is, 4 patients that were not diagnosed at the time of analysis, 6 patients with low- to high-grade dysplasia (site: larynx), 5 patients with malignant tumours at other sites and 10 patients with benign tumours at other sites. Hence, we analysed for this study the samples of 62 well-defined subjects: 22 with HNSCC (site: larynx and pharynx), 21 with benign tumours (site: larynx and pharynx: vocal cord polyps, nodules or intra-cordal cysts) and 19 healthy controls.

Head and neck cancer is classified according to primary tumour stage (T), regional lymph node stage (N) and distant metastasis stage (M). The 22 HNSCC patients in this study included 9 patients with early-stage disease (T $\leqslant$ 2; N = 0; M = 0), 11 patients with late-stage disease (all higher TNM classifications) and 2 patients were not staged.

Table 1 summarises the average clinical characteristics of the study groups, including tumour site and stage for the HNSCC patients. Detailed information on each volunteer, including age, gender, smoking/alcohol status, and TNM stage of the HNSCC patients is given in the Supplementary Information, in Tables S1–S3.

Considering that alcohol consumption is a well-known HNSCC risk factor, the percentage of HNSCC patients consuming alcohol (13%) in this study might seem unusually low. Note that this seeming abnormality resulted from the very low alcohol consumption in the Israeli general population: only 5.1% of the Israeli population reported frequent alcohol consumption during a recent national survey (Neumark *et al*, 2007). In this context it should also be noted that most patients in this cohort suffered from glottic carcinoma, which are more closely correlated to smoking than to alcohol consumption.

Ethical approval has been obtained from the institutional review board of Carmel Medical Center, and the study has been registered at http://clinicaltrials.gov. The treatment decisions were based solely on the conventional diagnosis described above.

Study design. The primary aim of this cross-sectional comparative study was to distinguish HNSCC patients with tumours in the larynx and pharynx from two control groups: (i) from patients with benign tumours at the same sites and (ii) from healthy (tumour-free) controls. The secondary aim was to distinguish subpopulations within the HNSCC group according to two different criteria: (iii) according to tumour location (larynx or pharynx) and (iv) according to tumour stage (early stage or advanced stage, see Supplementary Information, Table S1. The effect of two confounding factors was tested: (1) smoking and (2) gender. Conventional diagnosis served as reference standard.

This single-centre pilot study with a limited patient cohort of 62 (after application of exclusion criteria, see section Patients of Patients and Methods) was designed as a feasibility test of a nanomaterial-based breath test for HNSCC of the larynx and

Table 1. Descriptive statistics of the 62 volunteers included in this study					
Classification	Number of patients	Age (years)	Gender (M : F)	Smoking (%)	Alcohol use (%)
All samples					
HNSCC[a]	22	62 ± 12	19 : 3	59	13
benign tumours	21	55 ± 14	14 : 7	57	14
HNSCC stages					
Early-stage HNSCC[b]	9	67 ± 9	8 : 1	—	—
Late-stage HNSCC[b]	11	58 ± 13	9 : 2	—	—
Unknown stage	2	—	—	—	—
HNSCC site: larynx					
HNSCC	12	63 ± 11	12 : 0	—	—
Benign tumours	20	56 ± 14	13 : 7	—	—
HNSCC site: pharynx					
HNSCC	10	61 ± 13	7 : 3	—	—
Benign tumours	1	—	—	—	—
Healthy (tumour free)	19	50 ± 12	6 : 14	25	10

Abbreviation: HNSCC = squamous cell carcinoma of the head and neck.
[a]Head and neck cancer squamous cell carcinoma.
[b]See Supplementary Information, Table S1.

pharynx, with the aim of delivering a proof-of-concept that would justify a large-scale, multicentre trial.

Breath sample collection. Exhaled alveolar breath was collected in a controlled manner, as described in Peng *et al* (2009); Peng *et al* (2010) and Hakim *et al* (2011). The volunteers were usually sampled on the morning of the day scheduled for biopsy under general anaesthesia. None of the volunteers consumed food, tobacco or alcohol during a (overnight) 12-h interval before the breath collection. The breath sampling process is described in the Supplementary Information, section S2. Briefly, the inhaled air was cleared of ambient contaminants by a 3- to 5-min lung washout with filtered air. The subjects then exhaled through a separate exhalation port of the mouthpiece against 10- to 15-cm H_2O pressure to ensure closure of the vellum so that nasal entrainment of gas is excluded. Exhaled breath is a mixture of alveolar air and respiratory dead space air. In an automatic single-step process, the dead space was separated from the alveolar breath. Only the alveolar breath was used for analysis. Samples were collected in duplicates per test person for the analysis with GC–MS and with the nanomaterial-based sensors.

Characterisation of the breath samples. The breath samples were analysed using two separate, entirely independent approaches:

1. Chemical analysis of the breath samples with the aim of identifying the VOCs that show statistically different concentrations in the compared subpopulations, using GC–MS. External standards were measured for compound identification/quantification, according to the recommendation in Bajtarevic *et al* (2009); Ligor *et al* (2009); Sponring *et al* (2009) and Filipiak *et al* (2010). The GC–MS analysis is described in detail in Supplementary Information, section S3.
2. Analysis of the breath samples with an array of six nanomaterial-based sensors, combined with discriminant factor analysis (DFA) as statistical pattern recognition algorithm (the details of the statistical analysis are given in section Sensor array), with the aim of identifying and distinguishing specific patterns for HNSCC and non-malignant lesions in the larynx and pharynx. The array included five sensors based on gold

nanoparticles with different organic ligands and one sensor based on single-walled carbon nanotubes capped with an organic cap-layer (see references (Tisch and Haick, 2010a,b)). A detailed description of the sensor array and the constituent electronic nose setup can be found in Supplementary Information, sections S4 and S5. The breath analysis process with the sensors array setup is described in Supplementary Information, section S5.

Statistical analysis

Gas chromatography/mass spectrometry. Compounds showing significant differences (cut-off *P*-value: 0.05) between the studied subpopulations were determined from the GC–MS results, using non-parametric Wilcoxon/Kruskal–Wallis tests for data that cannot be assumed to be normally distributed (Wilkoxon, 1945). Statistical analysis was performed using JMP, version 9.0.0 (SAS Institute Inc., Cary, NC, USA, 1989–2005).

Sensors array. Each of the six sensors in the array responded to all (or to a certain subset) of the VOCs found in the exhaled breath samples. Specific patterns and predictive models for HNSCC and subgroups of HNSCC were derived from the sensor array output, using DFA (Ionescu *et al*, 2002). Discriminant factor analysis is a linear, supervised pattern recognition method for reducing the multidimensional experimental data, in which the classes to be discriminated are defined before the analysis is performed. Discriminant factor analysis was also used as a heuristic to select the best sensors for each application out of the repertoire of six, by filtering out non-contributing sensors. The reason for selecting a certain set of sensing features for a particular problem was directly derived from their ability to discriminate between the various classification groups. The input variables for DFA were the four features extracted from the sensors' time-dependent resistance responses (see Supplementary Information, sections S4 and S5). The four sensing features were related to the normalised resistance change at the beginning of the exposure, at the middle of the exposure and at the end of the exposure (with respect to the value

of sensors resistance in vacuum prior to the exposure), and to the area beneath the time-dependent resistance response, as described in Supplementary Information, section S5. Discriminant factor analysis determines the linear combinations of the input variables such that the variance within each class is minimised and the variance between classes is maximised. The DFA output variables (i.e., canonical variables (CVs)) are obtained in mutually orthogonal dimensions; the first CV is the most powerful discriminating dimension. The classification success was estimated through leave-one-out cross-validation in terms of the number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) predictions. Given n measurements, the model was computed using n-1 training vectors. The validation vector that was left out during the training phase was then projected onto the model, producing a classification result. All possibilities of leave-one-sample-out were considered, and the classification accuracy was estimated as the averaged performance over the n tests. Pattern recognition and data classification were conducted using MATLAB (The MathWorks).

RESULTS

Chemical analysis of the breath samples by GC–MS. In total, 207 VOCs were identified in each individual breath sample, and 141 VOC were present in >80% of the breath samples. From these we excluded five compounds that have been identified previously from the GC–MS chromatograms of pristine Tenax material from unused ORBO 420 Tenax TA sorption tubes: methylene chloride, acetaldehyde, L-cysteine sulfonic acid, malonic acid and naphthalene (tentative identification by spectral library match; compounds library of the National Institute of Standards and Technology, Gaithersburg, MD 20899-1070 USA; see Supplementary Information, section S3; Amal *et al*, 2012). Overall, 30 VOCs had low similarity indices (<80) during spectral library match and were therefore excluded. Hence, 106 compounds were further analysed. A full list of these compounds is given in Supplementary Information, section S7. Since the null hypothesis for normal distribution of the GC–MS data was not fulfilled for these compounds, according to Shapiro–Wilk tests, we used non-parametric Wilcoxon/Kruskal–Wallis tests (cut-off value: $P=0.05$) for the comparative analysis of the GC–MS data. We compared: (i) HNSCC (22) to tumour-free healthy states (19); (ii) HNSCC (22) to benign tumour states (21); and (iii) benign tumour states (21) to tumour-free healthy states (19). The GC–MS data of the subgroups within the HNSCC group was not analysed because of the small sample size. Three VOCs were found to be of statistical significance for the separation of the groups. They were tentatively identified through spectral library match as ethanol, 2-propenenitrile, and undecane (see Table 2). The compound identity was confirmed through retention time match after the measurement of high purity external standards for these three compounds (obtained from Sigma-Aldrich, Israel; see Supplementary Information, section S3) – a procedure that has been recommended by several researchers in the field (Bajtarevic *et al*, 2009; Ligor *et al*, 2009; Sponring *et al*, 2009; Filipiak *et al*, 2010).

The concentrations of ethanol, 2-propenenitrile and undecane were significantly elevated in the breath of HNSCC patients, as compared to the healthy controls (see Table 2). Only undecane was significantly elevated for benign tumour states, as compared to the healthy controls, while the increase of ethanol and 2-propenenitrile was subsignificant. No compounds showed significant differences in concentration between the breath samples of patients with HNSCC and benign tumours.

Sensors array studies of the breath samples of HNSCC patients and controls. In this study, we have tested the feasibility of the

Table 2. Volatile organic compounds from exhaled breath samples that significantly distinguished (P<0.05) between HNSCC and healthy states, HNSCC and benign tumours, and benign tumours and healthy states

| Compound | CAS no. | Chemical group | Retention time (min) | m/z | Compound concentration (p.p.b.) | | | | | | | | P-values (Wilcoxon non-parametric method) | | Relatively elevated in |
| | | | | | HNSCC | | Benign tumours | | Healthy | | Room air (unfiltered) | | | | |
					Median	Inter-quartile	Median	Inter-quartile	Median	Inter-quartile	Median	Inter-quartile	HNSCC vs healthy	Benign tumours vs healthy	
Ethanol	64-17-5	Alcohol	2.52	31	1198.3	2848.1	1077.0	1950.3	504.8	698.7	2118.5	1536.2	0.050		HNSCC
2-Propanenitrile	107-13-1	Nitrile	3.047	53	11.9	4.4	11.1	6.4	9.5	2.5	11.1	8.3	0.038		HNSCC
Undecane	1120-21-4	Alkane	25.8	57	219.7	308.3	162.4	365.8	30.2	124.5	218.1	396.8	0.044	0.043	HNSCC/benign tumours

Abbreviations: CAS = chemical abstracts service; HNSCC = squamous cell carcinoma of the head and neck; m/z = mass-to-charge ratio (main mass).

sensors array approach for identifying HNSCC in the larynx and pharynx regions and for distinguishing these malignancies from either tumour-free healthy states or from benign tumours at the same sites. For this purpose we have developed several predictive models based on breath samples collected from 62 volunteers, using DFA. The first model distinguished 22 HNSCC patients from 19 healthy controls with a sensitivity of 77% and a specificity of 90% after cross-validation, resulting in an overall classification accuracy of 83% (see Table 3). Figure 1A shows the DFA plot obtained from the responses of three non-correlated sensing features from the same sensor (organic functionality: 1-decanethiol; see Supplementary Information, Table S4). The HNSCC and healthy states formed two well-defined clusters along the first CV (CV1; see Supplementary Information, section S4), with little overlap and only few misclassified samples.

The second model separated the 22 HNSCC patients from the 21 benign tumour patients with the same classification success (sensitivity: 77%, specificity: 90%, overall classification accuracy: 83%; see Table 3). Figure 1B shows the DFA plot of the second model, obtained from the responses of three sensing features from two different sensors, among them the separating sensor from model 1 (see Supplementary Information, Table S4). The CV1 clusters of the HNSCC and benign tumour states were well separated showing little overlap.

Figure 1C shows the third DFA model that was developed for separating the benign tumours from the healthy states, using three sensing features from two sensors, among them again one of the main separating sensor from models 1 and 2 (see Supplementary Information, Table S4). Classification accuracy after cross-validation was 73% (see Table 3).

With the fourth model we attempted distinguishing between HNSCC at the two different sites: larynx and pharynx. Excellent and stable separation could be achieved between the two groups, using three sensing features from three different sensors (see Supplementary Information, Table S4). The classification success during cross-validation was also excellent, with values for sensitivity, specificity and overall accuracy of over 90% for a combination of three sensing features (see Table 3). Figure 1D shows the excellent separation between the HNSCC larynx and HNSCC pharynx clusters along the CV1 axis.

We have derived a fifth model for staging HNSCC. For this purpose we have divided the HNSCC patients into two subgroups comprising of 9 patients with early-stage disease ($T \leqslant 2$; $N = 0$; $M = 0$) and 11 patients with late-stage disease (all higher classifications). Supplementary Information, Table S1 lists the TNM classifications of all HNSCC patients and their division into early- and late-stage disease. The classification success was even

higher than for the cancer site classification in model 4, with an overall classification accuracy of 95% for a combination of three features (see Table 3 and Supplementary Information, Table S4). Figure 1E shows the clearly separated and non-overlapping clusters of early and late-stage disease.

DISCUSSION

Characteristic VOC profiles of HNSCC. The study of exhaled VOCs that could serve as markers of cancer is an active field of research and > 1000 compounds have been reported during the past two decades. A summary and discussion of the most important cancer-related VOCs (several hundreds), together with their associated biochemical pathways, can be found in a recent review (Haick *et al*, 2014).

In this study, we have observed statistically different concentration profiles of ethanol, 2-propenenitrile and undecane for the three groups of interest: HNSCC (in the larynx and pharynx) patients, patients with benign tumours at the same sites and healthy controls. The median concentrations of all three compounds increased monotonically from healthy states over benign tumours to HNSCC, that is, conc. (healthy states) < conc. (benign tumours) < conc. (HNSCC; see Table 2). It seems that biochemical cellular processes associated with tumour formation lead to an increase of these three compounds, whereby the processes associated with malignancy seem to further increase the concentration levels. The increase in concentration was significant for ethanol, 2-propenenitrile and undecane in the HNSCC breath samples, as compared to the healthy states. In contrast, only undecane showed a significant increase in concentration for the benign tumour states, as compared to the healthy states, whereas the increase in the ethanol and 2-propenenitrile concentrations was subsignificant. The direct comparison between HNSCC states and benign tumour states showed that the three compounds were increased only subsignificantly for HNSCC. Nevertheless, HNSCC and benign tumours in the larynx and pharynx regions could in principle be identified unambiguously via the direct comparison with the healthy controls, using their concentration profiles of ethanol, 2-propenenitrile and undecane (see Table 2). We would like to note that VOCs serving as cancer markers do not necessarily increase, as observed in this study. Indeed, tumour-specific metabolic alteration due to cancer also would predict that certain compounds decrease.

Ethanol is one of the four major endogenous volatiles in exhaled breath, besides methanol, acetone, and isoprene (Jones, 1985). In agreement to our observations, ethanol has recently been identified

Table 3. Predictive models based on the sensor array output for the pairwise comparison between HNSCC, benign tumours and healthy states, as well as between different sites and stages of HNSCC

DFA model no.	Target group	Control group	Total number of samples	Number of sensing features for DFA model	Accuracy (%; after cross-validation)	Sensitivity (%)	Specificity (%)
1	HNSCC	Healthy	41	3	83	77	90
2	HNSCC	Benign tumours	43	3	84	77	90
3	Benign tumours	Healthy	40	3	73	71	75
4	HNSCC: larynx	HNC: pharynx	22	3	91	92	90
5	Early HNSCC	Late HNSCC	21	3	95	100	91

Abbreviations: DFA = discriminant factor analysis; HNC = head and neck cancer; HNSCC = squamous cell carcinoma of the head and neck. The samples were classified using DFA, keeping the ratio between the number of independent measurements and the number of DFA input features below 7. Sensitivity, specificity and overall classification accuracy were calculated via leave-one-out cross-validation.
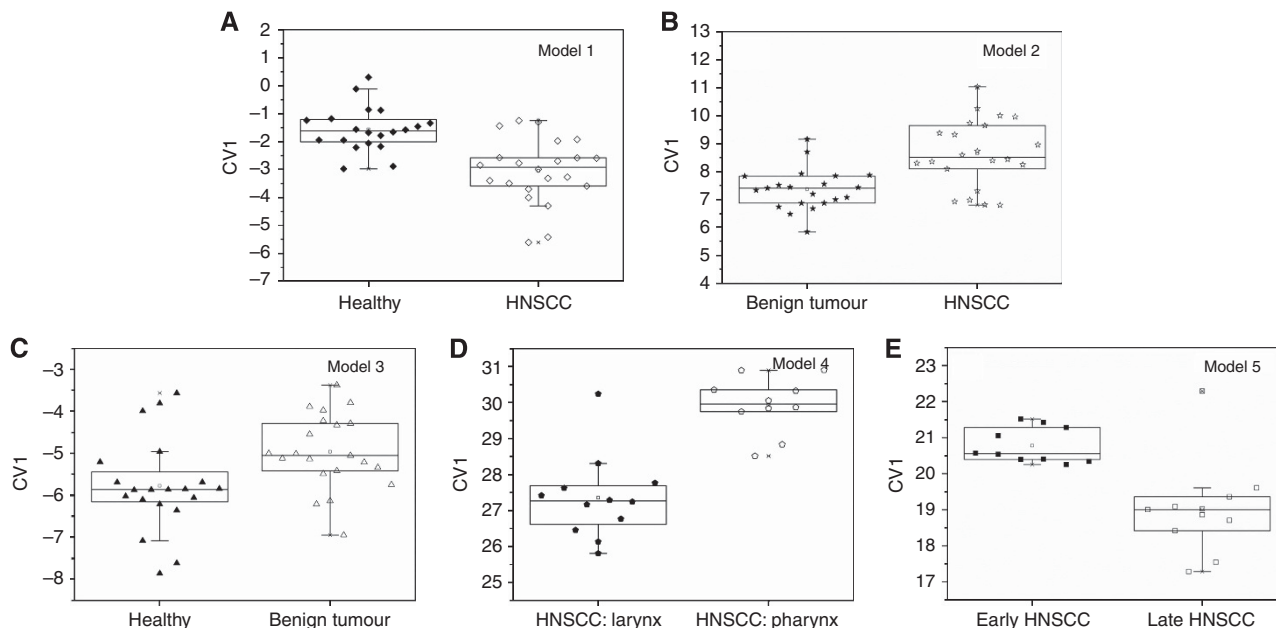
Figure 1. Predictive DFA models for distinguishing: (**A**) HNSCC patients from healthy (tumour-free) subjects, (**B**) HNSCC from benign tumour patients, (**C**) benign tumour patients from healthy subjects, (**D**) larynx malignancy from pharynx malignancy and (**E**) early HNSCC from late HNSCC.

as potential marker for laryngeal malignancy by Grácia *et al* (2014). Considering that alcohol abuse is an important confounding factor for HNSCC, we have carefully verified that the elevated concentrations were not due to the drinking habits of the individual participants (see Supplementary Information, Tables S1–S3). Table 1 shows that the three study populations of HNSCC, benign tumours and healthy controls are well matched with respect of the drinking habits of the participants: only a small percentage of the sampled subjects (13%, 14% and 10%, respectively, for HNSCC, benign tumours and healthy controls) consumed alcohol regularly. Note that the unusually low alcohol consumption of the HNSCC patients in this study is related to the very low alcohol consumption among the general population in Israel, and is not due to a selection bias during patient recruitment (Neumark *et al*, 2007). It should also be noted that most patients in this cohort suffered glottic carcinoma, which are more closely correlated to smoking than to alcohol consumption.

2-Propenenitrile (acrylonitrile) can be found as an environmental pollutant in cigarettes and in car exhaust, and was classified as a class 2B carcinogen (possibly carcinogenic) by The International Agency for Research on Cancer (IARC, 1999). In this study, the percentage of smokers among the HNSCC- and benign tumour patients (59% and 57%, respectively, see Table 1) was much higher than among the healthy controls (25%). It is reasonable to assume that this difference has contributed to the observed significant elevation of the 2-propenenitrile levels in HNSCC patients: this compound could indeed have been accumulated in the body of the patients due to continued cigarette exposure. Smoking is considered a major risk factor for HNSCC. Hence, we propose considering an elevated 2-propenenitrile level in the breath that stems from cigarettes and/or car exhaust as an *exogenous (risk) marker* of HNSCC. Interestingly, 2-propenenitrile has also been found in elevated concentrations in the breath samples of gastric cancer patients, as compared to healthy controls (Amal *et al*, 2013).

The elevated levels of undecane, an alkane that typically occurs in the breath of healthy subjects (Haick *et al*, 2014), could result from oxidative stress: alkanes are produced in the body mainly due to peroxidation of PUFA, known as lipid peroxidation by ROS (Kneepkens *et al*, 1994), which could be involved in the

pathogenesis of HNSCC. Oxidative stress could also play a role in the formation of benign tumours. The possible biochemical pathways of HNSCC require a larger study with larger and more diverse patient population.

The VOC profiles observed here for a well-defined group of HNSCC patients with tumours at two particular sites (larynx and pharynx) are different from the ones in our previously published pilot study (Hakim *et al*, 2011). Our earlier study had included a heterogeneous group of HNC patients with a variety of histologically different tumours at different sites (including larynx, pharynx, skin, salivary gland, etc). The differences in the patient groups, as well as the differences in the applied experimental methods (e.g., sample pre-concentration, GC–MS instrumentation and measurement protocol) have most probably contributed to the different VOC profiles.

Grácia *et al* (2014) has reported seven VOCs as possible markers of laryngeal carcinoma. Among these, only ethanol was observed also in this study. Again, this disagreement can be easily understood in terms of the pronounced differences in the patient groups and in the applied experimental methods (e.g., sample collection, pre-concentration, GC–MS instrumentation and measurement protocol) between the two studies. Although their observations are interesting, one should bear in mind that the sample size in this pilot study is rather small, and that the observed trends should be confirmed through a larger clinical trial before drawing far-reaching conclusions. Furthermore, it should be noted that the (unfiltered) room-air levels of all three distinguishing compounds were in the order of magnitude of the concentrations in the breath samples, and in the case of ethanol even slightly exceeded the breath concentrations (see Table 2). We would like to mention that all volunteers performed a lung washout prior to the breath collection by inhaling for 3- to 5-min through a filter cartridge, thus greatly reducing the concentration of exogenous VOCs and removing 99.99% of the exogenous compounds from the air during inspiration (see Supplementary Information, section S2). Consequently, we did not observe any correlation between the room-air VOC concentrations on a particular day and the concentration levels in the breath samples that were collected on the same day. However, we cannot altogether exclude the

possibility that the results were confounded through previous inhalation or up-take of the three VOCs from the hospital environment, storage in the body and subsequent gradual expiration. In this case, the concentration in the exhaled breath might be correlated with the period of previous exposure, rather than with the disease state.

Breath-prints of HNSCC (malignancy, site and staging). Although the chemical analysis of the breath samples by GC–MS has provided interesting information for biochemical research, and has shown that the VOC profiles of HNSCC, benign tumours and healthy states are distinctly different, the sensor array approach would be more practical as future diagnostic tool.

Our results showed that excellent separation could be achieved between HNSCC and healthy states, as well as HNSCC and benign tumours (see Figures 1A and B). The separation was achieved using only 3 sensing features out of a pool of 24 features, for a total of over 40 independent measurements in both models (see Table 3). The low feature : sample ratio effectively reduced the risk of over-fitting the experimental data. Cross-validation yielded stable classification accuracies of over 80% for both problems (see Table 3). In contrast, the benign tumours were less well-separated from the healthy states and the clusters overlapped more prominently (Figure 1C). The classification accuracy after cross-validation reached only just over 70% (see Table 3). Only insignificant improvement could be achieved by increasing the number of sensing features for the DFA model to four (data not shown). This might indicate that the breath-prints of the benign tumours were indeed closer to the healthy states than the malignant tumours, in agreement with the GC–MS results.

We have also demonstrated the feasibility of the breath test for determining the location and stage of HNSCC tumours. The malignant tumours of the larynx and pharynx had similar histology (squamous cell carcinoma, SCC). Hence it is reasonable to expect that part of their characteristic VOCs would be similar. With this in mind we had grouped them together in the first three predictive models (see Table 3). However, cancers at the two sites also display important differences: different risk factors, different prognosis and different treatment schemes. Therefore one can expect that part of their emitted VOCs would be different. The DFA analysis performed in this study has confirmed this hypothesis. For clinical use, the identification of the HNSCC tumour site would not be very relevant, because tumours of the pharynx can easily be visibly accessed. However, the demonstrated ability might be interesting for future use in other SCCs. The successful demonstration of HNSCC staging with the sensor array is both of scientific and clinical interest, since prognosis and therapy approach critically depend on this factor.

Smoking, alcohol abuse, HPV status, gender (HNSCC is more common in men than in women) and advanced age are the major risk factors for HNSCC. Since all these have their own characteristic VOC profiles, they could easily confound the sensor-based breath-prints of HNSCC (Amann *et al*, 2007; Amann *et al*, 2010; Tisch and Haick, 2010a). It is therefore essential to test the stability of the derived predictive models against these risk and/or potential confounding factors. Table 1 shows that the healthy controls were strongly mismatched both with the HNSCC and with the benign tumour patients in regard of their smoking status and gender ratio. Note that the strong gender mismatch is a result of the recruitment of healthy controls among the patient's accompanying persons, which were usually their spouses. Since benign and malignant lesions of the head and neck occur predominantly in men, the patient groups were male-dominated, while the healthy control group was female-dominated. This gave us the opportunity to test the stability of the first and the third predictive model (HNSCC *vs* health and benign tumours *vs* healthy, respectively, see Table 3) against smoking and gender.

We could not test the stability against HPV status, since all subjects in this study were tested and confirmed to be HPV negative. Also, the three test groups were well matched with regard to their alcohol habits: only just over 10% of subjects in each group drank regularly (see Table 1). Again, we would like to remind the reader that the unusually low alcohol consumption among the HNSCC patients in this study stems from the low alcohol consumption of the general population at the collection locality (Neumark *et al*, 2007). The age distribution within the groups was relatively similar and the differences between the groups were subsignificant ($P < 0.05$). Also, we could not test the effect of the confounding factors on the second predictive model distinguishing between HNSCC and benign tumours, since the two groups were well matched with regard to all the recorded confounding factors.

In model 1, the HNSCC group included 86% men and 59% smokers, as opposed to 30% men and 25% smokers among the healthy controls. Trying to separate (i) men and women and (ii) smokers and non-smokers, first in the healthy group alone, then in the HNSCC group alone, yielded in all the cases overlapping clusters and arbitrary classification accuracy ($\sim 50\%$) during leave-one-out cross-validation. Model 3, which separated the benign tumour patients (67% men, 57% smokers) from the healthy controls, was also found to be stable against smoking status and gender, using the same strategy as before.

Possible future relevance for clinical practice. Upper aero-digestive tract lesions are prevalent among the general population, but only few of these lesions are malignant. Benign and malignant conditions may present with similar clinical symptoms, so that upper aero-digestive tract endoscopy with biopsy and histopathological tissue analysis are essential for diagnosis.

A future nanomaterial-based breath test for the simultaneous detection of malignant and benign head and neck conditions would be ideally suited to precede and complement conventional invasive procedures. Breath testing is fast, simple and non-invasive. Hence, the test would be suitable for identifying at-risk individuals that should undergo further investigations. In this setup, breath testing could indicate malignancy prior to the endoscopic examination, thus allowing a well-directed, systematic search for malignant lesions, including hidden and small lesions that could otherwise be missed. Other possible future applications could include mass screening of high-risk populations such as heavy smokers.

The results of the breath test could potentially provide valuable complementary information for distinguishing malignant and benign lesions with similar appearance although it will not make tissue sampling obsolete.

Conclusion. In summary, we have presented a pilot study that demonstrates the feasibility of VOC profiling for identifying HNSCC. Three compounds were identified as possible markers for HNSCC and/or benign lesions of the head and neck: ethanol, 2-propenenitrile and undecane. Possible biochemical explanations for the increase of these compounds in HNSCC patients were discussed, but the clinical usefulness of these VOC profiles for patient classification would be limited. As more realistic basis for a future point-of-care HNSCC test, we have demonstrated the discriminative power of a nanomaterial-based sensor array. Breath-prints derived from the sensory output provided allowed distinguishing HNSCC from benign tumours and from healthy states, as well as identifying site and stage of HNSCC, irrespective of gender and tobacco consumption, which constitute important confounding/risk factors for HNSCC. In the long term, the presented results could eventually lead to the development of a simple breath test that may be used to aid and complement conventional HNSCC diagnosis. In the short term, this single-centre pilot study with a limited patient cohort has delivered a proof-of-concept for the proposed nanomaterial-based breath

test that would justify a large-scale, multicentre trial with a more realistic ratio of malignant to non-malignant head and neck conditions.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Amal H, Ding L, Liu BB, Tisch U, Xu ZQ, Shi DY, Zhao Y, Chen J, Sun RX, Liu H (2012) The scent fingerprint of hepatocarcinoma: in-vitro metastasis prediction with volatile organic compounds (VOCs). *Int J Nanomedicine* **7**: 4135–4146.

Amal H, Leja M, Broza YY, Tisch U, Funka K, Liepniece-Karele I, Skapars R, Xu ZQ, Liu H, Haick H (2013) Geographical variation in the exhaled volatile organic compounds of gastric cancer. *J Breath Res* **7**: 047102.

Amann A, Miekisch W, Pleil J, Risby T, Schubert J (2010) Methodological issues of sample collection and analysis of exhaled breath. *European Respiratory Society Monograph* **49**: 96–114.

Amann A, Spanel P, Smith D (2007) Breath analysis: the approach towards clinical applications. *Mini Rev Med Chem* **7**(2): 115–129.

Argiris A, Karamouzis MV, Raben D, Ferris RL (2008) Head and neck cancer. *Lancet* **371**(9625): 1695–1709.

Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, Ligor T, Filipiak W, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Haidenberger A, Buszewski B, Miekisch W, Schubert J, Amann A (2009) Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC Cancer* **9**: 348.

Barash O, Peled N, Hirsch FR, Haick H (2009) Sniffing the unique 'odor print' of non-small-cell lung cancer with gold nanoparticles. *Small* **5**(22): 2618–2624.

Broza YY, Haick H (2013) Nanomaterial-based sensors for detection of disease by volatile organic compounds. *Nanomedicine* **8**(5): 785–806.

Filipiak W, Sponring A, Filipiak A, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J (2010) TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells in-vitro. *Cancer Epidemiol Biomarkers Prev* **19**: 182–195.

García RA, Morales V, Martín S, Vilches E, Toledano A (2014) Volatile organic compounds analysis in breath air in healthy volunteers and patients suffering epidermoid laryngeal carcinomas. *Chromatographia* **77**: 501–509.

Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, Oneill HJ (1985) Volatile organic-compounds in exhaled air from patients with lung-cancer. *Clin Chem* **31**: 1278–1282.

Haddad RI, Shin DM (2008) Recent advances in head and neck cancer. *N Engl J Med* **359**(11): 1143–1154.

Haick H, Broza YY, Mochalski P, Ruzsanyi V, Amann A (2014) Assessment, origin, and implementation of breath volatile cancer markers. *Chem Soc Rev* **43**: 1423–1449.

Hakim M, Billan S, Tisch U, Peng G, Dvrokind I, Marom O, Abdah-Bortnyak R, Kuten A, Haick H (2011) Diagnosis of head-and-neck cancer from exhaled breath. *Br J Cancer* **104**(10): 1649–1655.

IARC (1999) Re-evaluation of some organic chemicals, hydrazine and hydrogen peroxide. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Vol. 71, pp. 43-108. World Health Organization Distribution and Sales: Geneva.

Ionescu R, Broza YY, Shaltieli H, Sadeh D, Zilberman Y, Feng X, Glass-Marmor L, Lejbkowicz I, Müllen K, Miller A, Haick H (2011) Detection of multiple sclerosis from exhaled breath using bilayers of polycyclic aromatic hydrocarbons and single-wall carbon nanotubes. *ACS Chem Neurosci* **2**(12): 687–693.

Ionescu R, Llobet E, Vilanova X, Brezmes J, Sueiras JE, Calderer J, Correig X (2002) Quantitative analysis of NO2 in the presence of CO using a single tungsten oxide semiconductor sensor and dynamic signal processing. *Analyst* **127**: 1237–1246.

Jones AW (1985) Excretion of low-molecular weight volatile substances in human breath: focus on endogenous ethanol. *J Anal Toxicol* **9**(6): 246–250.

Kneepkens CM, Lepage G, Roy CC (1994) The potential of the hydrocarbon breath test as a measure of lipid peroxidation. *Free Radic Biol Med* **17**(2): 127–160.

Konvalina G, Haick H (2014) Sensors for breath testing: from nanomaterials to comprehensive disease detection. *Acc Chem Res* **47**(1): 66–76.

Leunis N, Boumans ML, Kremer B, Din S, Stobberingh E, Kessels AG, Kross KW (2013) Application of an electronic nose in the diagnosis of head and neck cancer. *Laryngoscope* **124**(6): 1377–1381.

Ligor M, Ligor T, Bajtarevic A, Ager C, Pienz M, Klieber M, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Buszewski B, Miekisch W, Schubert J, Amann A (2009) Determination of volatile organic compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry. *Clin Chem Lab Med* **47**(5): 550–560.

Marur S, D'Souza G, Westra WH, Forastiere AA (2010) HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol* **11**: 781–789.

Mendis S, Sobotka PA, Euler DE (1994) Pentane and isoprene in expired air from humans: gas-chromatographic analysis of single breath. *Clin Chem* **40**: 1485–1488.

Miekisch W, Schubert JK, Noeldge-Schomburg GFE (2004) Diagnostic potential of breath analysis—focus on volatile organic compounds. *Clin Chim Acta* **347**: 25–39.

Neumark YD, Lopez-Quintero C, Grinshpoon A, Levinson D (2007) Alcohol drinking patterns and prevalence of alcohol-abuse and dependence in the Israel National Health Survey. *Isr J Psychiatry Relat Sci* **44**(2): 126–135.

O'Neill HJ, Gordon SM, O'Neill MH, Gibbons RD, Szidon JP (1988) A computerized classification technique for screening for the presence of breath biomarkers in lung cancer. *Clin Chem* **34**: 1613–1618.

Pai SI, Westra WH (2009) Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu Rev Pathol* **4**: 49–70.

Peng G, Tisch U, Adams O, Hakim M, Shehada N, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, Haick H (2009) Diagnosing lung cancer in exhaled breath using gold nanoparticles. *Nat Nanotechnol* **4**(10): 669–673.

Peng S, Hakim M, Broza Y, Billan S, Abdah- Brotnyak R, Kuten A, Tisch U, Haick H (2010) Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. *Br J Cancer* **103**: 542–551.

Phillips M, Gleeson K, Hughes JMB, Greenberg J, Cataneo RN, Baker L, McVay WP (1999) Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. *Lancet* **353**: 1930–1933.

Phillips M, Greenberg J, Awad J (1994) Metabolic and environmental origins of volatile organic compounds in breath. *J Clin Pathol* **47**: 1052–1053.

Schmutzhard J, Rieder J, Deibl M, Schwentner IM, Schmid S, Lirk P, Abraham I, Gunkel AR (2008) Pilot study: volatile organic compounds as a diagnostic marker for head and neck tumors. *Head Neck* **30**(6): 743–749.

Shuster G, Gallimidi Z, Heyman-Reiss A, Dovgolevsky E, Billan S, Abdah-Bortnyak R, Kuten A, Engel A, Shiban A, Tisch U, Haick H (2010) Classification of breast cancer precursors through exhaled breath. *Breast Cancer Res Treat* **126**: 791–796.

Song G, Qin T, Liu H, Xu G, Pan YY, Xiong FX, Gu KS, Sun GP, Chen ZD (2010) Quantitative breath analysis of volatile organic compounds of lung cancer patients. *Lung Cancer* **67**(2): 227–231.

Sponring A, Filipiak W, Mikoviny T, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J (2009) Release of volatile organic compounds from the lung cancer cell line NCI-H2087 in vitro. *Anticancer Res* **29**: 419–426.

Tisch U, Aluf Y, Ionescu R, Nakhleh M, Bassal R, Axelrod N, Robertman D, Tessler Y, Finberg JPM, Haick H (2012) Detection of asymptomatic nigrostriatal dopaminergic lesion in rats by exhaled air analysis using carbon nanotube sensors. *ACS Chem Neurosci* **3**(3): 161–166.

Tisch U, Haick H (2010a) Arrays of chemisensitive monolayer-capped metallic nanoparticles for diagnostic breath testing. *Rev Chem Eng* **26**: 171–179.

Tisch U, Haick H (2010b) Sensors based on monolayer-capped metal nanoparticles. In *Chemical Sensors*, Schwank J, Korotcenkov G (eds), Vol. 2: Nanostructured Materials, pp 141–202. Momentum Press: New York.

Wilkoxon F (1945) Individual comparisons by ranking methods. *Biometrics Bull* **1**: 80–83.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)