Title

Breath carbonyl levels in a human population of seven hundred participants

Journal

Journal of Breath Research, 14(4)

ISSN

1752-7155

Authors

McCartney, Mitchell M

Thompson, Carina J

Klein, Lauren R

et al.

Publication Date

2020-10-01

DOI

10.1088/1752-7163/ab8865

Peer reviewed

HHS Public Access

Author manuscript

J Breath Res. Author manuscript; available in PMC 2022 December 21.

Published in final edited form as:

J Breath Res. ; 14(4): 046005. doi:10.1088/1752-7163/ab8865.

Breath carbonyl levels in a human population of seven hundred participants

Mitchell M. McCartney[1], **Carina J. Thompson**[1], **Lauren R. Klein**[1], **Josephine H. Ngo**[1], **Jacqueline D. Seibel**[1], **Fauna Fabia**[1], **Leslie A. Simms**[1], **Eva Borras**[1], **Brian S. Young**[2], **Juven Lara**[2], **Michael W. Turnlund**[2], **Anh P. Nguyen**[3,4,5], **Nicholas J. Kenyon**[3,4,6], **Cristina E. Davis**[1,*]

[1]Mechanical and Aerospace Engineering, University of California Davis, Davis, CA, USA

[2]Revelar (Pulse Health LLC.), 2611 SW 3rd Ave. Suite 216, Lake Oswego, OR, USA

[3]Department of Internal Medicine, 4150 V Street, Suite 3400, University of California Davis, Sacramento, CA, USA

[4]VA Northern California Health Care System, 10535 Hospital Way, Mather, CA, USA

[5]Division of Rheumatology, Allergy and Clinical Immunology, 451 Health Sciences Drive, University of California Davis, Sacramento, CA, USA

[6]Center for Comparative Respiratory Biology and Medicine, University of California Davis, CA, USA

Abstract

Oxidative stress is associated with numerous health conditions and disorders, and aldehydes are known biomarkers of oxidative stress that can be non-invasively measured in exhaled human breath. Few studies report breath aldehyde levels in human populations, and none claim participant numbers in the hundreds or more. Further, the breath community must first define the existing aldehyde concentration variance in a normal population to understand when these levels are significantly perturbed by exogenous stressors or health conditions. In this study, we collected breath samples from 692 participants and quantified C_4–C_{10} straight chain aldehyde levels. C_9 aldehyde was the most abundant in breath, followed by C_6. C_4 and C_5 appear to have bimodal distributions. *Post hoc*, we mined our dataset for other breath carbonyls captured by our assay, which involves elution of breath samples onto a solid phase extraction cartridge, derivatization and liquid chromatography-quadrupole time of flight mass spectrometry (LC-qTOF). We found a total of 21 additional derivatized compounds. Using self-reported demographic factors from our participants, we found no correlation between these breath carbonyls and age, gender, body mass index (BMI), ethnicity or smoking habit (tobacco and marijuana). This work was preceded by a small confounders study, which was intended to refine our breath collection procedure. We found that breath aldehyde levels can be affected by participants' using scented hygiene products such as lotions and mouthwashes, while collecting consecutive breath samples, rinsing the mouth with water, and filtering inspired air did not have an effect. Using these parameters to guide our

* cedavis@ucdavis.edu .

sampling, subjects were instructed to avoid the prior conditions to providing a breath sample for our study.

Keywords

aldehydes; carbonyls; exhaled breath; human population study

1. Introduction

Oxidative stress is involved in many disorders, such as attention-deficit/hyperactivity disorder (ADHD)[1], Parkinson's Disease[2], autism[3] and vitiligo[4]. Although not the cause, diabetics also suffer increased levels of oxidative stress[5]. Inflammatory responses to airway diseases, especially asthma and COPD, are heavily affected by oxidative stress.[6] Monitoring oxidative stress levels has the potential to aid treatments of many ailments, and follow patient response to therapy and wellness routines.

Due to an imbalance with antioxidant defenses, oxidative stress produces reactive oxygen species (ROS). As targets of ROS, lipids are oxidized into polyunsaturated fatty acids (PUFAs). In turn, PUFAs undergo secondary reactions and can be converted into carbonyls, such as aldehydes.[7] Thus, researchers have demonstrated that aldehydes can serve as biomarkers for oxidative stress.[8] The aliphatic aldehydes have high enough gas-liquid Henry constants to diffuse readily into expired air.

Human breath provides a non-invasive way to measure health. Researchers, including our group, have well established that human breath contains hundreds of metabolites.[9, 10] Breath metabolites can serve as biomarkers of rhinovirus infection[11], diabetes[12] and COPD.[13] Multiple studies confirm that aldehydes and other carbonyls in breath correlate to lung cancers[14–19] Furthermore, human bronchial epithelial cell cultures have been shown to serve as models of a portion of the airway system,[20] and these cells emit volatile organic compounds (VOCs), including aldehydes, when undergoing oxidative stress.[21]

For breath biomarkers to serve as diagnostic or health monitoring tools, we must understand what normal or typical concentrations exist in the human population. Once the variance has been defined for a general population, breath researchers can more readily understand when exogenous factors or health conditions perturb oxidative stress biomarkers to unusual levels. To date, few large-scale studies have been conducted that report quantitative concentrations of metabolites in human breath. In 1999, a study of 50 people listed quantified breath volatiles, but in concentrations relative to an internal standard.[22] Blanchet et al. presented an impressive multivariate analyses of breath samples from 1,417 adults related to endogenous and exogenous confounding factors, but did not report quantitative levels of the analytes.[23] A final example analyzed breath from 21,582 persons, providing quantitative statistics but only on a single compound: breath ethanol.[24]

In this work, we aim to move breath research forward with a large-scale study that quantifies straight-chain, C_4–C_{10} breath aldehyde concentrations in a human population. We summarize aldehyde levels from 692 participants. Further, we attempt to correlate breath

samples to demographic factors of the participants, such as age, BMI, ethnicity and smoking habits. We also include results from a small confounders study that refined our sample collection protocol. Carbonyls from breath were extracted, derivatized and quantified via liquid chromatography-quadrupole time of flight (LC-qTOF) mass spectrometry.

2. Materials and method

This study was approved by the UC Davis IRB for collection and analysis of human breath samples, IRB# 1158345. All subjects signed an informed consent prior to participating.

2.1 Breath collection, chemical analysis

Multiple sets of experiments were performed in this study, but all breath samples were collected, processed and chemically analyzed in the following manner.

2.1.1 Breath collection.—Breath samples were collected by having participants exhale into commercial 10 L Tedlar bags (Part 1219–7000-GD, Environmental Sampling Supply). Bags were conditioned under vacuum at 80 °C for at least 4 h prior to sample collection. Participants did not eat or apply any scented product for at least 2 h prior to sample collection. Participants remained in their setting for 10 min to allow their lungs to equilibrate to the background air. They were instructed to respire normally to maintain tidal breathing, exhaling through a valve to fill the bags, after which the valve was closed. The majority of breath samples were processed within 2 h of collection and all were processed within 16 h.

2.1.2 Metabolite extraction.—Solid phase extraction (SPE) was used to extract carbonyls from the breath samples. SPE cartridges contained 300 mg of unbonded silica sorbent in a 6 mL cartridge (Part CUSIL136, UCT Inc.). Cartridges were preconditioned with 5 mL of 0.6% acetic acid in methanol, then 5 mL of methanol, followed by applying 60 psi of ultra-high purity nitrogen for 30 min to one end. We confirmed this was enough time to completely dry the cartridges, which we tested by weight (data not shown).

The port on the 10 L bag containing the breath sample was connected directly to an SPE cartridge via a rubber stopper. With the breath sample connected on one end, the opposite end of the cartridge was connected to a house vacuum line that extracted the breath sample at a rate of 3.310 SLPM ± 0.050. This flow rate was pre-determined and did not present issues of aldehyde breakthrough in the cartridge sorbent (data not shown). The breath sample valve was opened and the sample was left to extract until the bag was completely empty.

Once the breath sample was completely loaded onto the cartridge, the cartridge was eluted with 1 mL of a solution of 0.6% acetic acid, 40% methanol and 59.4% water. Two internal standards were added into the eluate: 50 μL of 650 nM dodecanal, used to qualitatively confirm derivatization and 50 μL of 650 nM TAMRA-C_{16}-amide ($C_{46}H_{64}N_4O_5$), a pre-derivatized amide with 100% conversion that was quantitatively used to normalize data. With every batch of breath samples, a conditioned SPE with no loaded breath sample was also eluted, providing a “cartridge/reaction blank” to ensure that the residual aldehyde concentration from the silica, solvent and reagents remained very low.

2.1.3 Carbonyl derivatization.—Derivatization was performed to stabilize carbonyls in solution. The derivatization agent, 6-AO-TAMRA ($C_{32}H_{37}N_5O_6$) is a fluorophore that contains an alkoxyamine functional group, which combines with carbonyls to form a stable oxime bond.[25] Fluorescence detectors have been used to measure these derivatized carbonyls, but herein we used mass spectrometry. This solution was reacted with 150 μL of the derivatization agent (650 nM 6'-AO-TAMRA) and 50 μL of a 60 μM 5-methoxyantranilic acid (catalyst), which was mixed, lidded and left to react at room temperature for 60 min. The reaction was stopped with the 100 μL addition of 1 M NH_4CO_3 with a pH adjusted to 10.0. Samples were kept at −20 °C until analysis.

2.1.4 LC-qTOF analysis.—Chemical analysis was performed on an Agilent 1290 Infinity LC and 6545 qTOF system. Samples were chilled on the sample tray at 8 °C. 5 μL of the derivatized breath sample was injected into the system. We used a pre-column, XTERRA MS C18 5 μm, 2.1 × 5 mm (Part #186007896, Waters), connected to an XTERRA MS C18 5 μm, 2.1 × 150 mm (Part #186000454, Waters) LC column for carbonyl separation. The column compartment was set to 35 °C. The flow was set to 0.2 mL/min throughout. Solvent A consisted of water with 0.18% formic acid, 10 mM ammonium acetate, pH adjusted to 4.0 with ammonium hydroxide. Solvent B was methanol. The solvent gradient was: 0 min 40% B, 24 min 82% B, 26 min 100% B, 32 min 100% B, 33 min 40% B. The qTOF was set to positive mode. The sheath gas flow, nozzle voltage, drying gas flow, nebulizer, capillary voltage and fragmentor voltage were all optimized (Supplemental Figure 1). The mass spectrometer scanned from 100 to 1700 m/z at a rate of 3 spectra/sec. A "system blank" (no sample injection) occurred every 20 injections to ensure the LC-qTOF system remained free of contamination. Known concentrations of derivatized aldehydes were injected alongside samples to generate the calibration curves required for quantitative analysis of C_4-C_{10} aldehydes.

2.1.5 Data processing, analysis.—LC-qTOF data were processed using MassHunter Quantitative Analysis (Version B.07, Agilent Technologies). Masses of derivatized, ionized carbonyls were extracted with a window of ± 20 ppm and peak areas were integrated. Peak areas were normalized to the internal standard. Statistical analyses were performed in MATLAB (Version R2017a, MathWorks) and PLS_Toolbox (Version 8.6, Eigenvector Research Inc.). Before multivariate analyses, data underwent Pareto scaling. A p-value of $p \leq 0.05$ was used throughout for statistical significance tests.

2.2 Confounders study

A confounders study was conducted to see if certain activities and products would alter breath aldehyde levels. These included: consecutive breath samples, rinsing mouth with water prior to breath sample, filtering inspired air and using scented/flavored toiletries.

For consecutive breath samples, participants provided two breath samples with a 5 min rest between. For the mouth rinse, participants rinsed their mouths with ~20 mL of water for 10 s just prior to providing a breath sample.

To filter inspired air, participants inhaled through an activated carbon filter (Product 6005, 3M). They then exhaled into the breath bag, repeating this process until the bag was full.

Three products were tested to see if their usage could alter breath aldehyde levels. A hand lotion, a lip balm and a mouthwash were confirmed to contain aldehydes in the following manner: an enhanced lipid matrix removal kid (QuEChERS, Agilent Technologies) was used to prepare 40 g of the lip balm and lotion. One milliliter of each subsequent solution was derivatized per normal sample procedures. For the mouthwash, 5 mL was loaded directly onto the SPE cartridge, which underwent normal elution and subsequent derivatization.

Participants (3 male, 3 female) were asked to provide a breath sample, use one of the products, then provide a second breath sample to see if breath aldehyde levels were elevated by the products. Each participant tested each product 3 times.

2.4 Human population study

Participants were recruited in the Davis and Sacramento regions of California. Sampling stations were set up at both indoor and outdoor locations, such as farmers markets, sporting events and inside office buildings, with permissions acquired as necessary. For the human population study, participants were asked to fill out a one page survey. They self reported age, gender, height, weight, ethnicity, smoking habits (tobacco and marijuana), known illnesses/diseases, any medications taken within 2 weeks, and the area where they live (development type, zip code, number of blocks from a freeway). No further confirmation was conducted on their claims; participants underwent no health screening. Participants were provided a $5 gift card as compensation for their time.

3. Results and discussion

We share the results of a large-scale study to quantify C_4-C_{10} straight chain aldehyde (SCA) levels in a human population. We began with a confounders study to refine a sample collection strategy.

3.1 Confounders study

We tested if certain activities or products would influence breath aldehyde measurements, targeting C_4-C_{10} straight chain aldehydes. A summary of these findings is provided (Figure 1).

3.1.1 Consecutive breath samples.—We first investigated whether providing two consecutive breath samples would alter breath aldehyde levels, which had implications on the experimental design of the rest of the confounders study. Ten participants provided 1–3 paired breath samples for a total of n=19 before/after comparisons. C_4-C_9 SCA concentrations increased modestly in the 2nd breath sample (Figure 1A), with the only significant increase being C_5 ($p=0.043$), which had an average increase of 0.010 ± 0.030 ppb. C_{10} SCA had a slight decrease, on average, (-0.001 ± 0.029 ppb) in the 2nd consecutive sample. We concluded that providing two consecutive samples did not have, in general, a major significant effect on SCA breath levels. With these results, we designed the rest of the confounders study such that participants would be their own control and provide a before/after paired sample: a participant would provide an initial breath sample, then perform an activity and finally provide a second breath sample.

3.1.2 Water rinse.—Six participants provided 13 pre- and post-mouth rinse samples. Per a paired t-test, no aldehyde concentration was significantly affected by a mouth rinse, with average changes ranging from −0.019 to 0.017 ppb for C_4-C_{10} SCAs (Figure 1B). These results were not surprising, given the hydrophobic properties of these straight chain aldehydes. We suspect a mouth rinse might help to dislodge food particles that might contain aldehydes, yet in this study participants were already required to not have eaten within 2 h of sample collection. Furthermore, these results were similar to the consecutive breath sample study and essentially corroborate the finding that consecutive breath samples do not affect measurements.

3.1.3 Filtering inspired air.—While it has been well established that aldehydes are constituents of human breath, aldehydes also naturally exist in the environment. Thus, it might be possible that inspired, exogenous compounds could confound measurements of respired, endogenous breath constituents, although this was examined in a study of C_3-C_9 breath aldehydes. In their study, Poli et al. observed that aldehydes were more abundant in breath than in background air.[15] Still, should the risk exist, one way to prevent potential contamination is to filter inspired air, which typically can be accomplished by having the volunteer inhale through a sorbent-filled mouthpiece, reducing or eliminating background VOCs. For our measurements, the participant would then directly exhale into our breath collection bags.

We compared 50 unfiltered and 50 filtered breath samples from 18 participants (Figure 1C). While a paired t-test found these differences to be significant for C_4-C_9, the median reduction of aldehyde concentrations ranged only from 0.005 to 0.018 ppb. These values were similar to the changes seen from consecutive breath samples and the mouth rinse, despite most of those not having a statistically significant difference. Our conclusion is that the filter used did not provide a substantial decrease in the breath aldehyde measurement to necessitate its use.

3.1.4 Using common toiletries.—We directly analyzed three toiletries and confirmed they contained at least one aldehyde (Table 1). Paired t-tests were used to determine if a person’s breath aldehyde levels were altered after using these aldehyde-containing products.

The hand lotion contained C_7-C_{10} SCAs. All four compounds were significantly higher in breath samples after participants applied the lotion, with mean increases of 0.006, 0.334, 0.063 and 0.144 ppb, respectively (Figure 1D). However, there was no significant difference when participants applied the lip balm, despite it containing C_6-C_8 & C_{10} SCAs (Figure 1E). We noted that the lotion was strongly perfumed, and we suspect that it caused elevated breath samples simply by participants’ breathing in the lotion’s aroma and exhaling it into the sample bags. The lip balm did not have a potent smell, meaning the aldehydes may not have been as volatile or were in lower concentrations in the balm relative to the lotion, and did not produce the same elevating effect.

The use of a cinnamaldehyde-containing mouthwash dramatically increased breath cinnamaldehyde levels. While we did not convert the instrument response to concentration for cinnamaldehyde, peak areas on average increased by 11,846% in breath samples just

after using the mouthwash (n=18 before/after comparisons, data not shown). The derivatized cinnamaldehyde peak was not quite 3× the standard deviation of the background noise levels, typically considered the threshold for the limit of detection, yet became very apparent after participants used the mouthwash. This experiment was repeated with two participants, who provided a breath sample just before mouthwash, just after mouthwash, then every 30 min for 2 h (Figure 1F). The mouthwash had a significant increase just after participants used it, but the breath levels returned to initial values within 30 min.

3.1.5 Confounders study conclusion.—

The purpose of this small confounders study was to demonstrate that breath readings can be altered by actions prior to breath sampling. We observed that toiletries can significantly elevate aldehyde levels in breath when these products contain aldehydes. Ideally, participants would be screened for products known to contain aldehydes, but this is an impractical task given the sheer number of products commercially available. The results of this study helped guide our final sample collection protocol for our larger human population study, which was refined to the following.

Participants:

- Could not eat, drink or smoke within 2 h prior to sampling
- Could not have applied any lotions, perfumes, lip balms, mouthwash, etc. within 2 h prior to sampling
- Must have sat in the same room or area for 5 min to allow lungs to equilibrate to environment

3.2 Human population study

Breath samples were collected from 692 participants across the Davis and Sacramento areas of California. Demographic statistics are presented (Table 2).

3.2.1 Concentration distributions of aldehydes.—

Provided are the distributions of breath aldehydes as measured in this human population, both tabulated (Table 3) and illustrated (Figure 2). Generally, aldehyde concentrations were less than 1 ppb.

Nonanal was the most abundant in human breath, with a mean value of 0.60 ppb, followed by hexanal at 0.318 ppb. These two aldehydes are the products of ω3 and ω6 fatty acid peroxidation, and the fatty acids are found in cell membrane phospholipids. Thus, it is not surprising for these to be the most abundant of the aldehydes measured. Butanal and pentanal present a bimodal distribution (Figure 2), both with a less frequent mode at lower concentrations. We currently do not have an explanation for this phenomenon.

There has been a large disagreement between reported concentrations of aldehydes in human breath among researchers (Table 3). Most published values have aldehydes less than or around single digit part-per-billion, but some studies have seen level as high as 24.5 ppb.[18] Still, the cited publications in Table 3 all compare healthy controls to lung cancer patients, and in all studies, every breath aldehyde had a higher mean or maximum among cancer patients. No participant included in this study reported having lung cancer but no conformations of any kind were made for self-reported illnesses/diseases. It may be

possible that our population sees lower values compared to other researchers' since we did not include lung cancer patients.

It is worth mentioning that several studies have reported breath aldehyde levels as measured in exhaled breath condensate (EBC).[13, 26] We would expect that concentrations in condensate would not be similar to the concentrations measured in exhaled gas, which is what we report in this study. Thus, we did not compare our measured concentrations to those reported in EBC.

3.2.2 Additional carbonyls.—After completion of this study, we mined this dataset *post hoc* for other carbonyls that were captured by this assay, derivatized by 6'-AO-TAMRA and measured by our LC-qTOF. This procedure was executed by creating a theoretical library of 84 carbonyls that have been reported in breath literature. The library contained saturated aldehydes and ketones, as well as unsaturated (alkenals, hydroxy-alkenals, etc.). Confirmation with chemical standards was done when possible. A total of 21 derivatized carbonyls were included in breath measurements (Table 4). These compounds were included in the subsequent demographic analysis.

3.3 Demographic correlations to breath aldehyde levels

The 692 participants also completed a one-page demographic survey. We attempted to correlate their breath samples to these self-reported factors. Chemometric analysis was conducted using the normalized peak area of the 21 detected carbonyls relative to the internal standard of each sample. Table 2 shows a summary of the demographic values reported by our volunteers.

3.3.1 Age.—Researchers have taken varying approaches to correlate age and breath VOCs. In some studies, samples were binned into decadal groups (e.g. ages 20–29, 30–39, etc.) and researchers looked for differences among these groups. Doing so, a study has found one breath aldehyde, dodecanal, decreases with age.[23] Dodecanal was not included in our study, as it was used as an internal standard to qualitatively verify derivatization of the sample.

Overall, our samples were skewed towards younger participants, a result of recruiting mostly within a college campus. We created a partial least squares (PLS) regression model using a random 66% of breath samples. PLS is a multivariate analysis that, in this case, attempts to correlate age to breath metabolite abundances. This model was then used to predict the age of the remaining 34% of samples and had very poor prediction capabilities (Figure 3A). The model did not predict anybody was higher than 42 years old; however, the oldest participant in the test set was 81. The difference in predicted and actual age correlated with actual age, as 100% of breath samples from participants aged ≤32 were predicted to be older than they were, whereas 94% of breath samples from participants >32 were under predicted.

We grouped samples into bins 10 years wide and performed a MANOVA. A canonicals variates (CV) plot (Figure 4A) shows that breath samples could not be separated according to their age based on our measured breath metabolites.

3.3.2 BMI.—Schwarz et al. looked at breath acetone levels and BMI, finding no correlation between the two.[27] Blanchet et al. reported 15 breath VOCs that tracked with BMI, but none were aldehydes.[23]

Similarly as with age, we both created a regression model and attempted to find differences by binning participants into four categories of BMI (underweight <18.5, normal 18.5–24.9, overweight 25–29.9, obese >30). The PLS model (Figure 3B), built with a random two-thirds of samples, was extremely ineffective at predicting the BMI of the remaining test set. While the test set included actual BMIs that spanned from 17–37.9, the predicted BMIs were 25.0 ± 0.2.

A MANOVA was performed on samples binned by BMI group (Figure 4B). The CV plot shows strong similarities between groups. The underweight group had high variance, as demonstrated by its 95% confidence ellipse, which is likely due to the relatively small number (n=22) of breath samples in this category.

3.3.3 Gender.—There have been few examples of breath metabolites with gender specificity, although none include the metabolites studied here. In a study of 216 participants, breath isoprene levels were higher in men than women by about 24 ppbv[28], although another isoprene study found no differences between genders.[29] In a small study of 47 volunteers, female and male samples separated by 11 breath VOCs, including the branched aldehyde 3-methylbutanal.[30] Blanchet et al. found several compounds with different abundances between genders in a set of 1,336 samples, although the difference was not strong enough for a random forest model to differentiate genders.[23]

Similar to Blanchet et al.[23], we were unable to build a discriminate gender model based on our measured breath metabolites (Figure 3C). Building a PLS-DA model from a random 2/3rds of breath samples and using it to predict the gender of the remaining 1/3rd resulted in a sensitivity of 55.1% and a specificity of 55.3%. The model was no better than randomly guessing the gender of the participants. A canonical variates (CV) plot from a MANOVA showed strong overlaps between genders (Figure 4C).

A 2-sample t-test found only one metabolite, benzaldehyde, with abundance differences between genders ($p<0.001$). On average, benzaldehyde was upregulated in men, with average abundances being 147% higher relative to women.

3.3.4 Ethnicity.—There have been almost no studies that compared breath profiles of different ethnic groups. One study found differences in nitric oxide levels in breath of White and South Asian children[31], but did not include carbonyls.

We did not observe ethnicity-based differences based on breath carbonyls. Our analysis was conducted on the ethnic groups from which we collected samples from at least 20 participants. This included four groups: White, Asian, Hispanic and Black. A PLS-DA model showed poor separation of ethnic groups based on breath samples (Figure 3D). An n-way analysis of variance did not find any compounds with significant differences among ethnic groups. A CV plot did not show separation by group (Figure 4D).

One might expect diet and culture to influence the breath metabolome. But our population, living in a well-integrated college campus community, likely enjoys a melting pot of cuisines and activities that do not shift metabolites in discrete ways. It might be suggested for future studies to compare samples from populations that have stronger cultural differences, such as between states or countries.

3.3.5 Smoking.—Others have investigated effects of cigarette smoking on breath VOCs. A small n=19 pilot study found increased levels of hydroxy-alkenals in the breath of smokers[32], but did not report on straight chain aldehydes. Two other studies have reported on breath compounds associated with smoking habits, but neither found associations with SCAs.[23, 33] A study of exhaled breath condensate from 12 smokers and 25 controls found elevated levels of hexanal and heptanal in the smokers.[34]

Studies on marijuana and breath have not yet correlated effects of marijuana smoking with endogenous metabolites, but instead have focused on the detection of Δ^9-tetrahydrocannabinol (THC) in breath.[35, 36]

Our study did not find correlations of measured breath aldehydes to our participants' self-reported smoking habits for neither tobacco nor marijuana (Figure 3, E and F). PLS-DA models could not classify participants by smoking habits. Separately for tobacco and marijuana, we built two PLS-DA models with a random 66% of the data sets and applied these models to the remaining 34% sets, or test sets. Prediction capabilities of smoking habits based on breath samples were very poor. Receiver operating characteristic curves for marijuana smokers had an area under the curve (AUC) of only 50.2%. The AUC for tobacco smokers was only 50.5%. MANOVA of each smoking type showed overlaps between the groups (Figure 4, E and F).

An n-way ANOVA found one compound, benzaldehyde, to correlate with tobacco use ($p<0.001$). Abundance of unk10 increased with tobacco use, with the average level in smokers being 1.8 times greater than non-smokers. Benzaldehyde has been shown to be upregulated in human airway epithelial cell cultures that have been exposed to oxidative stress agents.[21]

4. Conclusion

In this analysis of 692 samples, we quantify C_4–C_{10} straight chain aldehyde levels in human breath from a normal or "unremarkable" population. We feel this work is an especially important report for the breath community, as the literature contains sparse reports of participant numbers in the hundreds. Furthermore, almost no other report exists that quantifies breath compounds in a population this size. Our work reports what can be considered the typical or observed distributions of certain oxidative stress biomarkers, which serves as a baseline for future studies to investigate the exogenous and endogenous mechanisms that elicit significant changes to these observed levels. Aldehydes are known to correlate to oxidative stress, and this work presents an excellent starting point for future researchers to "move the dial" away from the normal levels reported herein, such

as comparison of healthy participants to people with confirmed afflictions, or through intervention studies such as changes to diet or exercise.

Mining of this dataset found a total of 21 breath carbonyls captured by our assay. Among these 21 compounds, we did not find strong correlations between measured breath carbonyls and self-reported demographic factors such as age, gender, ethnicity or smoking habits.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding was provided by Pulse Health, LLC. [CED, NJK]. This study was partially supported by: NIH award U01EB0220003-01 [CED, NJK]; the NIH National Center for Advancing Translational Sciences (NCATS) through grant UL1TR000002 [CED, NJK]; NIH award 1P30ES023513-01A1 [CED, NJK]; NIH award UG3-OD023365 [CED, NJK]. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

The authors gratefully acknowledge several people who provided indispensable intellectual and logistical guidance to this work: James Ingle, PhD; Linda McAllister MD, PhD; and Karen Copeland, PhD. We also thank the Food Safety and Measurement Facility at UC Davis, especially the Director of Research, Dr. Larry Lerno, for providing access to and maintaining the LC-qTOF. We also acknowledge Alexandria Falcon, Midori Huapaya-Renbarger, Sean Pham and Alexander Schmidt for their contributions to breath sample collection, who were all students at UC Davis during the time of this study. Finally, we once again thank the family members, friends and strangers who provided breath samples.

References

1. Joseph N; Zhang-James Y; Perl A; Faraone SV, Oxidative Stress and ADHD:A Meta-Analysis. Journal of Attention Disorders 2015, 19 (11), 915–924. [PubMed: 24232168]
2. Bosco DA; Fowler DM; Zhang Q; Nieva J; Powers ET; Wentworth P Jr; Lerner RA; Kelly JW, Elevated levels of oxidized cholesterol metabolites in Lewy body disease brains accelerate α-synuclein fibrilization. Nature Chemical Biology 2006, 2, 249. [PubMed: 16565714]
3. James SJ; Cutler P; Melnyk S; Jernigan S; Janak L; Gaylor DW; Neubrander JA, Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. Am J Clin Nutr 2004, 80 (6), 1611–1617. [PubMed: 15585776]
4. Arican O; Kurutas EB, Oxidative stress in the blood of patients with active localized vitiligo. Acta Dermatovenerol Alp Pannonica Adriat 2008, 17 (1), 12–6. [PubMed: 18454264]
5. Dierckx N; Horvath G; van Gils C; Vertommen J; van de Vliet J; De Leeuw I; Manuel-y-Keenoy B, Oxidative stress status in patients with diabetes mellitus: relationship to diet. Eur J Clin Nutr 2003, 57 (8), 999–1008. [PubMed: 12879095]
6. MacNee W, Oxidative stress and lung inflammation in airways disease. Eur J Pharmacol 2001, 429 (1–3), 195–207. [PubMed: 11698041]
7. Miekisch W; Schubert JK; Noeldge-Schomburg GF, Diagnostic potential of breath analysis--focus on volatile organic compounds. Clin Chim Acta 2004, 347 (1–2), 25–39. [PubMed: 15313139]
8. Del Rio D; Stewart AJ; Pellegrini N, A review of recent studies on malondialdehyde as toxic molecule and biological marker of oxidative stress. Nutr Metab Cardiovasc Dis 2005, 15 (4), 316–28. [PubMed: 16054557]
9. Amann A; Costello Bde L; Miekisch W; Schubert J; Buszewski B; Pleil J; Ratcliffe N; Risby T, The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. J Breath Res 2014, 8 (3), 034001.
10. Aksenov AA; Zamuruyev KO; Pasamontes A; Brown JF; Schivo M; Foutouhi S; Weimer BC; Kenyon NJ; Davis CE, Analytical methodologies for broad metabolite coverage of exhaled breath condensate. J Chromatogr B Analyt Technol Biomed Life Sci 2017, 1061–1062, 17–25.

11. Schivo M; Aksenov AA; Linderholm AL; McCartney MM; Simmons J; Harper RW; Davis CE, Volatile emanations from in vitro airway cells infected with human rhinovirus. J Breath Res 2014, 8 (3), 037110.
12. Minh Tdo C; Blake DR; Galassetti PR, The clinical potential of exhaled breath analysis for diabetes mellitus. Diabetes Res Clin Pract 2012, 97 (2), 195–205. [PubMed: 22410396]
13. Corradi M; Rubinstein I; Andreoli R; Manini P; Caglieri A; Poli D; Alinovi R; Mutti A, Aldehydes in exhaled breath condensate of patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2003, 167 (10), 1380–6. [PubMed: 12522029]
14. Callol-Sanchez L; Munoz-Lucas MA; Gomez-Martin O; Maldonado-Sanz JA; Civera-Tejuca C; Gutierrez-Ortega C; Rodriguez-Trigo G; Jareno-Esteban J, Observation of nonanoic acid and aldehydes in exhaled breath of patients with lung cancer. J Breath Res 2017, 11 (2), 026004.
15. Poli D; Goldoni M; Corradi M; Acampa O; Carbognani P; Internullo E; Casalini A; Mutti A, Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. J Chromatogr B Analyt Technol Biomed Life Sci 2010, 878 (27), 2643–51.
16. Schallschmidt K; Becker R; Jung C; Bremser W; Walles T; Neudecker J; Leschber G; Frese S; Nehls I, Comparison of volatile organic compounds from lung cancer patients and healthy controls-challenges and limitations of an observational study. J Breath Res 2016, 10 (4), 046007.
17. Buszewski B; Ligor T; Jezierski T; Wenda-Piesik A; Walczak M; Rudnicka J, Identification of volatile lung cancer markers by gas chromatography-mass spectrometry: comparison with discrimination by canines. Anal Bioanal Chem 2012, 404 (1), 141–6. [PubMed: 22660158]
18. Fuchs P; Loeseken C; Schubert JK; Miekisch W, Breath gas aldehydes as biomarkers of lung cancer. Int J Cancer 2010, 126 (11), 2663–70. [PubMed: 19839051]
19. Ulanowska A; Kowalkowski T; Trawinska E; Buszewski B, The application of statistical methods using VOCs to identify patients with lung cancer. J Breath Res 2011, 5 (4), 046008.
20. Yamaguchi MS; McCartney MM; Linderholm AL; Ebeler SE; Schivo M; Davis CE, Headspace sorptive extraction-gas chromatography-mass spectrometry method to measure volatile emissions from human airway cell cultures. J Chromatogr B Analyt Technol Biomed Life Sci 2018, 1090, 36–42.
21. Yamaguchi MS; McCartney MM; Falcon AK; Linderholm AL; Ebeler SE; Kenyon NJ; Harper RH; Schivo M; Davis CE, Modeling cellular metabolomic effects of oxidative stress impacts from hydrogen peroxide and cigarette smoke on human lung epithelial cells. J Breath Res 2019, 13 (3), 036014.
22. Phillips M; Herrera J; Krishnan S; Zain M; Greenberg J; Cataneo RN, Variation in volatile organic compounds in the breath of normal humans. Journal of Chromatography B 1999, 729 (1–2), 75–88.
23. Blanchet L; Smolinska A; Baranska A; Tigchelaar E; Swertz M; Zhernakova A; Dallinga JW; Wijmenga C; van Schooten FJ, Factors that influence the volatile organic compound content in human breath. J Breath Res 2017, 11 (1).
24. Gainsford AR; Fernando DM; Lea RA; Stowell AR, A large-scale study of the relationship between blood and breath alcohol concentrations in New Zealand drinking drivers. J Forensic Sci 2006, 51 (1), 173–8. [PubMed: 16423246]
25. Thomas G, Lara J, Noll C, Young B, Carlsen C, Mahon M, Ingle J. Methods, Systems, and Compositions for Detection of Aldehydes. 2017.
26. Corradi M; Folesani G; Andreoli R; Manini P; Bodini A; Piacentini G; Carraro S; Zanconato S; Baraldi E, Aldehydes and glutathione in exhaled breath condensate of children with asthma exacerbation. Am J Respir Crit Care Med 2003, 167 (3), 395–9. [PubMed: 12411284]
27. Schwarz K; Pizzini A; Arendacka B; Zerlauth K; Filipiak W; Schmid A; Dzien A; Neuner S; Lechleitner M; Scholl-Burgi S; Miekisch W; Schubert J; Unterkofler K; Witkovsky V; Gastl G; Amann A, Breath acetone-aspects of normal physiology related to age and gender as determined in a PTR-MS study. J Breath Res 2009, 3 (2), 027003.
28. 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 (3), 478–83. [PubMed: 16510318]

29. Kushch I; Arendacka B; Stolc S; Mochalski P; Filipiak W; Schwarz K; Schwentner L; Schmid A; Dzien A; Lechleitner M; Witkovsky V; Miekisch W; Schubert J; Unterkofler K; Amann A, Breath isoprene--aspects of normal physiology related to age, gender and cholesterol profile as determined in a proton transfer reaction mass spectrometry study. Clin Chem Lab Med 2008, 46 (7), 1011–8. [PubMed: 18605961]

30. Das MK; Bishwal SC; Das A; Dabral D; Varshney A; Badireddy VK; Nanda R, Investigation of gender-specific exhaled breath volatome in humans by GCxGC-TOF-MS. Anal Chem 2014, 86 (2), 1229–37. [PubMed: 24350824]

31. Sonnappa S; Bastardo CM; Stafler P; Bush A; Aurora P; Stocks J, Ethnic differences in fraction of exhaled nitric oxide and lung function in healthy young children. Chest 2011, 140 (5), 1325–1331. [PubMed: 21527509]

32. Gaugg MT; Gomez DG; Barrios-Collado C; Vidal-de-Miguel G; Kohler M; Zenobi R; Sinues PML, Expanding metabolite coverage of real-time breath analysis by coupling a universal secondary electrospray ionization source and high resolution mass spectrometry-a pilot study on tobacco smokers. J Breath Res 2016, 10 (1).

33. Peralbo-Molina A; Calderon-Santiago M; Jurado-Gamez B; Luque de Castro MD; Priego-Capote F, Exhaled breath condensate to discriminate individuals with different smoking habits by GC-TOF/MS. Sci Rep 2017, 7 (1), 1421. [PubMed: 28469199]

34. Andreoli R; Manini P; Corradi M; Mutti A; Niessen WM, Determination of patterns of biologically relevant aldehydes in exhaled breath condensate of healthy subjects by liquid chromatography/atmospheric chemical ionization tandem mass spectrometry. Rapid Commun Mass Spectrom 2003, 17 (7), 637–45. [PubMed: 12661015]

35. Manolis A; McBurney LJ; Bobbie BA, The detection of delta 9-tetrahydrocannabinol in the breath of human subjects. Clin Biochem 1983, 16 (4), 229–33. [PubMed: 6311456]

36. Coucke L; Massarini E; Ostijn Z; Beck O; Verstraete AG, Delta(9)-Tetrahydrocannabinol concentrations in exhaled breath and physiological effects following cannabis intake - A pilot study using illicit cannabis. Clin Biochem 2016, 49 (13–14), 1072–7. [PubMed: 27288550]

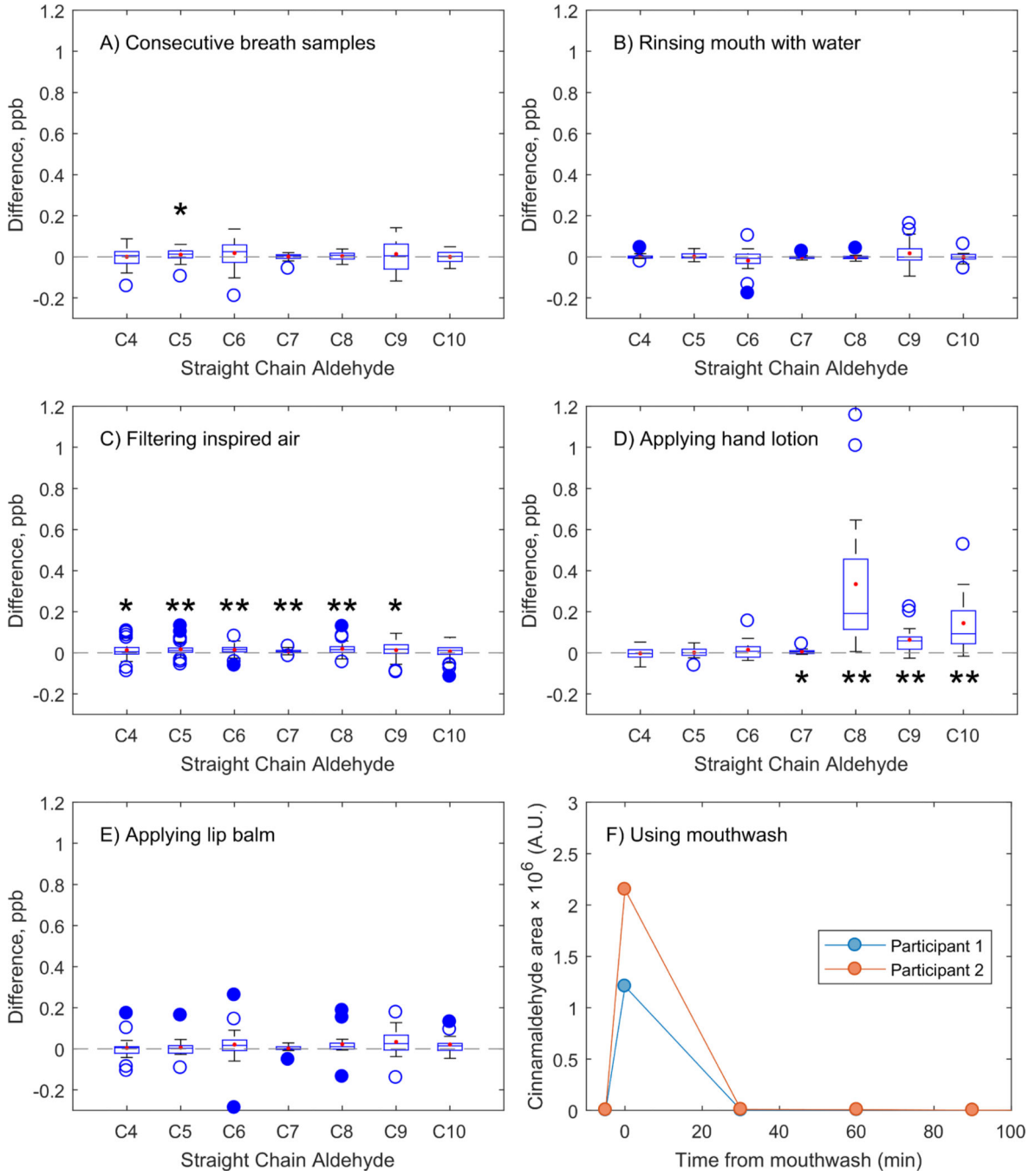

Figure 1:

Summary of confounders studies. See also: Table 1. A-E) Differences in breath aldehyde concentrations from paired before/after samples, with positive values indicating the confounder caused an increase. Asterisks indicate p-values from a paired t-test, with * for p>0.05 and ** for p>0.01. F) Immediately after using a mouthwash that contained cinnamaldehyde, the compound was found in breath but disappeared within 30 min of use.

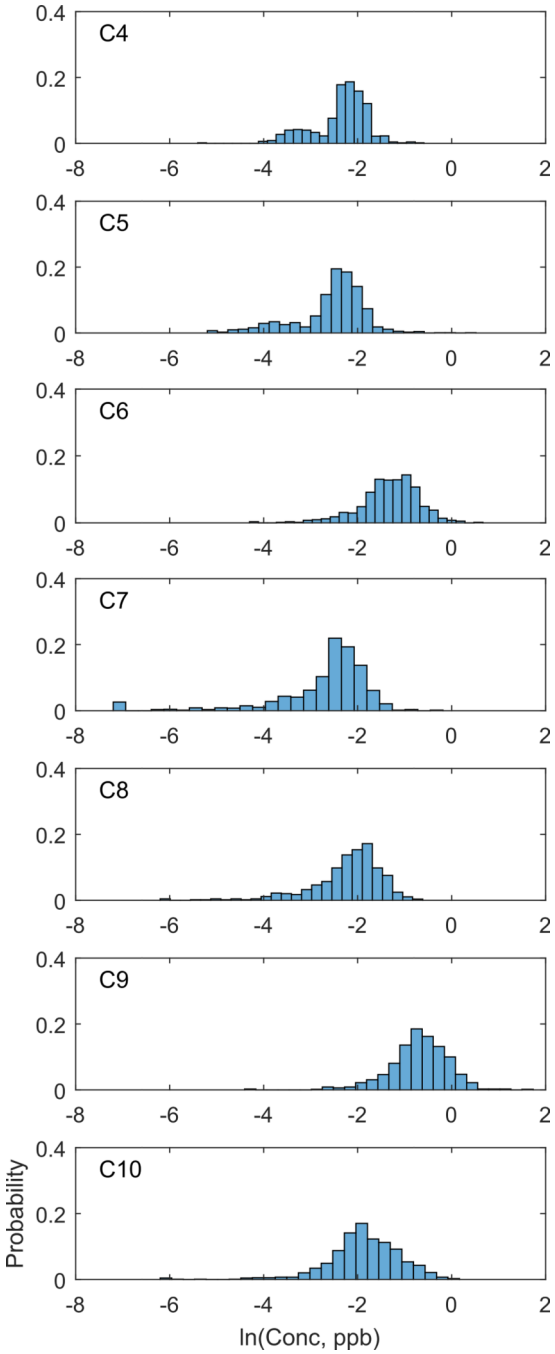

Figure 2:
Straight chain aldehyde concentration distributions. See also: Table 3.

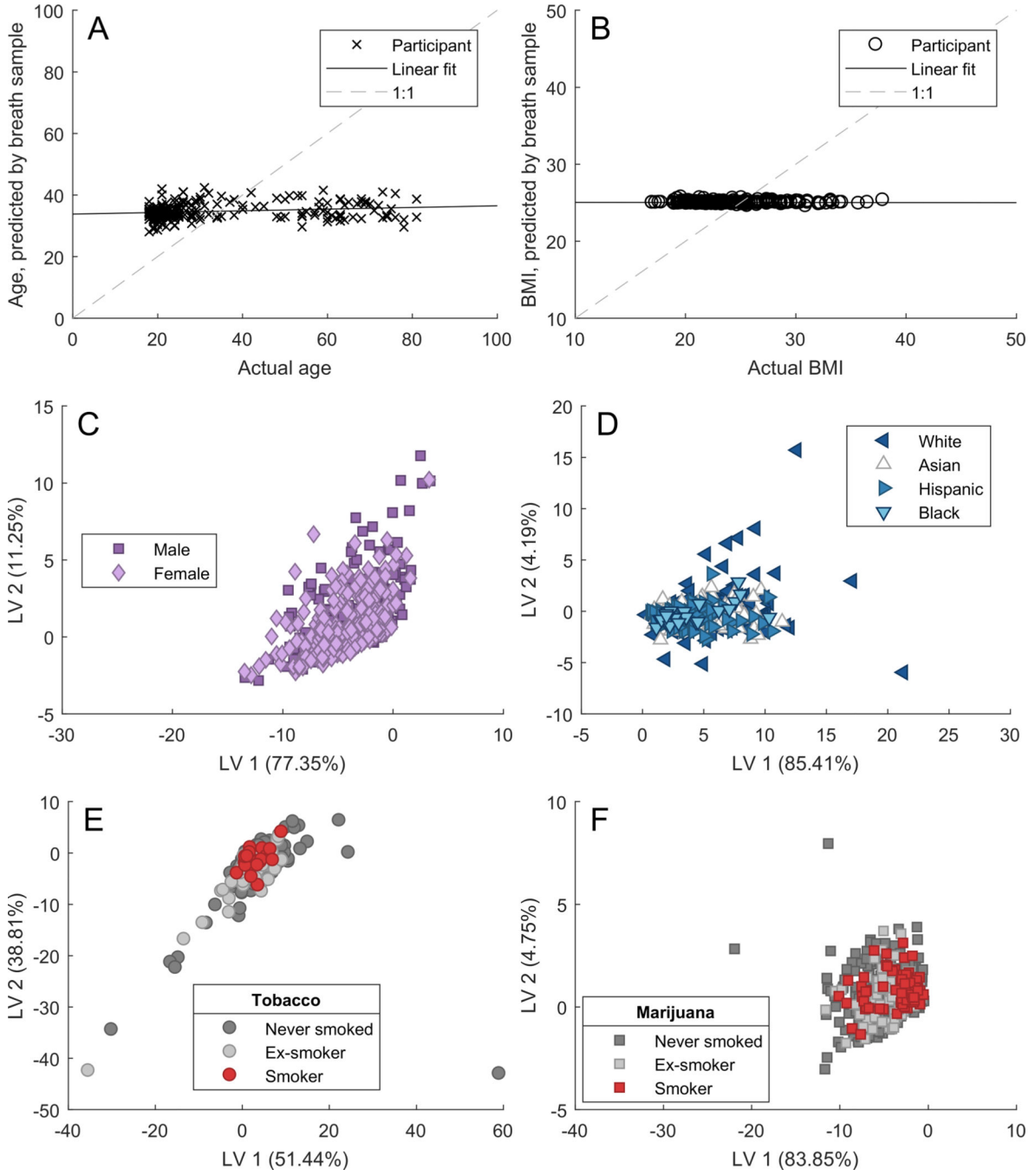

Figure 3:

No correlations were found between breath aldehyde levels and demographic factors. See also: Figure 4 A & B) For both age and BMI, two PLS regression models were built with a randomized 66% of samples. The remaining 34% of breath samples were inputted into these models to predict age and BMI. Both had poor prediction capabilities. C-F) Latent variable (LV) plots from PLS-DA analyses show no association between breath aldehyde levels and those demographic factors.

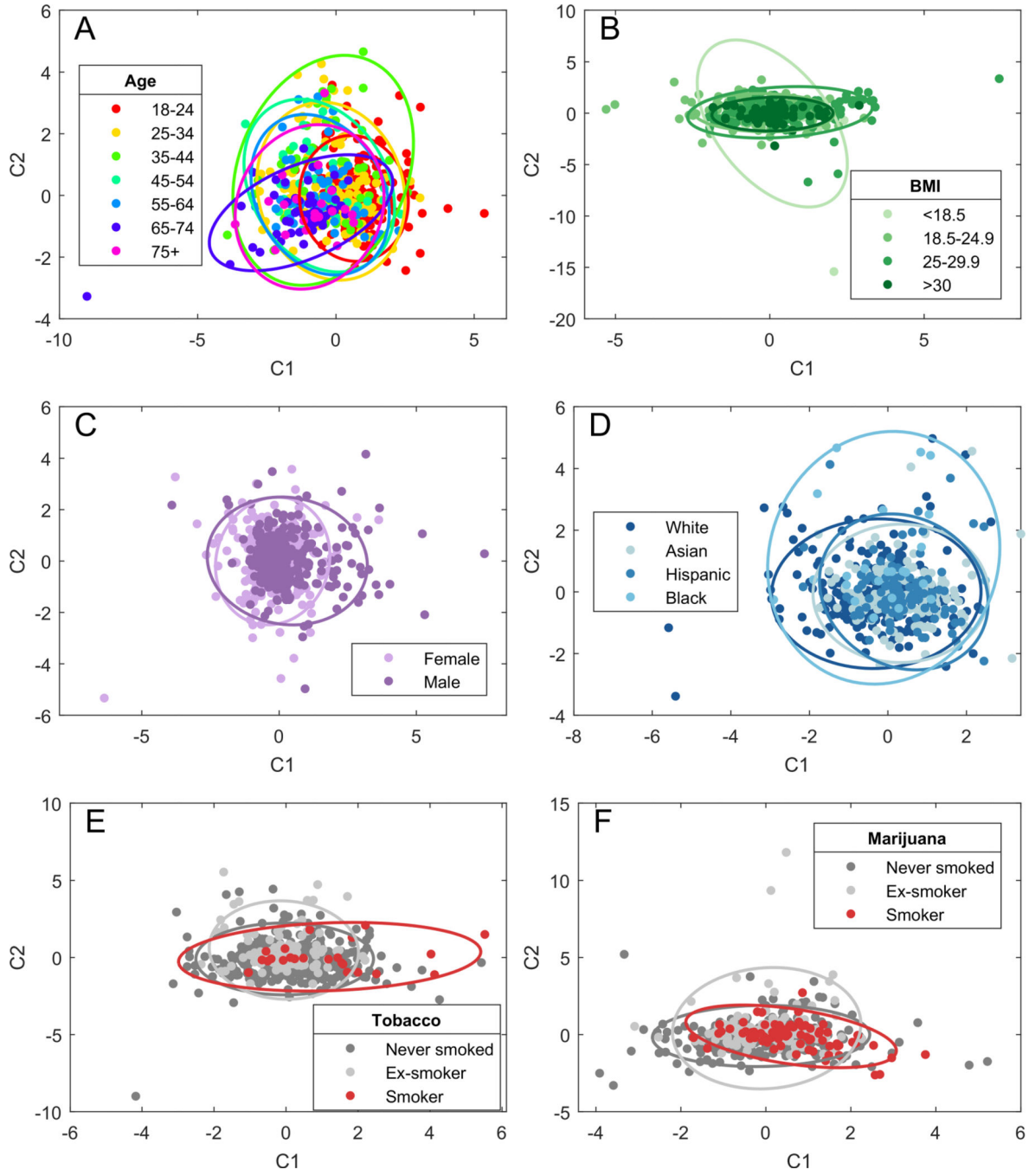

Figure 4:

Canonical variates score plots from MANOVA analyses for the impact of A) gender B) ethnicity C) age group D) BMI E) tobacco use and F) marijuana use on breath samples. Each point represents a breath sample from a participant, with the two canonical variates (C1 and C2) representing the most significant combinations of breath carbonyls related to the demographic groups. 95% confidence ellipses are shown.

Table 1:

Aldehydes were confirmed to be present in products used for the confounders study. Asterisks indicate that those aldehydes in breath samples were significantly elevated after using the product.

Product	Confirmed aldehydes present
Hand lotion, orange blossom scented	Heptanal * Octanal ** Nonanal ** Decanal **
Lip balm, orange ginger scented	Hexanal Heptanal Octanal Decanal
Mouthwash, “clean mint” flavored	Cinnamaldehyde **

* for $p \leq 0.05$

** for $p \leq 0.01$

Table 2:

Distribution of demographic factors (n=692 participants).

Factor	Distribution (%, or median ± IQR)
Age	26 ± 27 (minimum: 18)
BMI	23.8 ± 5.7
Gender	56% female, 44% male
Ethnicity	44.2% White 22.6% Asian 15.3% Hispanic, Latino or Spanish origin 10.2% Mixed Race 3.5% Black or African American 2.5% Middle Eastern or Northern African 0.9% Native Hawaiian or Other Pacific Islander 0.1% other
Tobacco, smoking habits	80.5% Never smoked 16.2% Ex-smoker 03.2% Smoker
Marijuana	68.5% Never smoked 19.3% Ex-smoker 12.1% Smoker

Table 3:

Distributions of straight chain breath aldehydes compared to the literature (Cit). See also: Figure 2.

Aldehyde	Study Results (ppb)	Literature data as published (ppb)	Cit.
butanal	MIN: 0.01 Q1: 0.08 MED: 0.11 MEAN: 0.11 Q3: 0.14 MAX: 0.55	 0.03–1.18 1.32–2.55 0.00–0.10 0.17–1.15	[16] [17] [18] [15]
pentanal	MIN: 0.01 Q1: 0.06 MED: 0.09 MEAN: 0.10 Q3: 0.12 MAX: 1.65	 0.03–3.10 0.00–0.78 0.00–7.20 0.11–1.23	[16] [18] [19] [15]
hexanal	MIN: 0.01 Q1: 0.20 MED: 0.28 MEAN: 0.32 Q3: 0.40 MAX: 1.91	 0.07–14.77 0.00–0.65 0.00–5.30 0.18–1.67	[16] [18] [19] [15]
heptanal	MIN: 0.00 Q1: 0.05 MED: 0.09 MEAN: 0.09 Q3: 0.12 MAX: 0.71	 0.00–0.65 0.10–0.60	[18] [15]
octanal	MIN: 0.00 Q1: 0.08 MED: 0.13 MEAN: 0.14 Q3: 0.17 MAX: 0.54	 0.02–0.70 0.03–2.18 0.18–0.86	[16] [18] [15]
nonanal	MIN: 0.01 Q1: 0.36 MED: 0.52 MEAN: 0.60 Q3: 0.73 MAX: 5.67	 0.03–2.98 0.38–12.4 0.18–1.80	[16] [18] [15]
decanal	MIN: 0.00 Q1: 0.11 MED: 0.16 MEAN: 0.20 Q3: 0.25 MAX: 1.16	 0.05–1.23 0.43–24.50	[16] [18]

Table 4:

Complete list of derivatized carbonyls captured and measured by our assay and used for demographic correlations.

#	Name	RT	Derivatized formula[†]	m/z (extracted)	Derivatized mass (measured)	Derivatized mass (exact)	Mass error (ppm)
1	unknown 1	4.944		644.2725	643.2647		
2	unknown 2	6.371		629.1515	628.1437		
3	formaldehyde[*]	7.654	C33H37N5O6	600.2831	599.2753	599.2744	1.54
4	acetaldehyde[*]	8.654	C34H39N5O6	614.2978	613.2900	613.2900	−0.05
5	unknown 3	9.504		614.2972	613.2894		
6	propanal[*]	9.954	C35H41N5O6	628.3143	627.3065	627.3057	1.31
7	unknown 4 (α,β-unsaturated C4 carbonyl)	11.538	C36H41N5O6	640.3133	639.3055	639.3057	−0.28
8	unknown 5 (saturated C4 ketone)	11.989	C36H43N5O6	642.3288	641.3210	641.3213	−0.51
9	butanal[*]	12.531	C36H43N5O6	642.3282	641.3204	641.3213	−1.45
10	benzaldehyde[*]	13.522	C39H41N5O6	676.3133	675.3055	675.3057	−0.27
11	pentanal[*]	14.591	C37H45N5O6	656.3437	655.3359	655.3370	−1.65
12	unknown 6 (saturated C5 ketone)	16.316	C37H45N5O6	656.3437	655.3359	655.3370	−1.65
13	hexanal[*]	16.657	C38H47N5O6	670.3609	669.3531	669.3526	0.70
14	heptanal[*]	18.602	C39H49N5O6	684.3762	683.3684	683.3683	0.18
15	octanal[*]	20.481	C40H51N5O6	698.3907	697.3829	697.3839	−1.48
16	nonanal[*]	22.273	C41H53N5O6	712.4081	711.4003	711.3996	1.01
17	decanal[*]	23.930	C42H55N5O6	726.4240	725.4162	725.4152	1.34
18	unknown 7	26.493		827.4672	826.4594		
19	undecanal[*]	26.708	C43H57N5O6	740.4380	739.4302	739.4309	−0.92
20	unknown 8	29.029		803.5434	802.5356		
21	unknown 9	29.315		623.3817	622.3739		

(*): compounds whose identities were confirmed by commercial standards.

(†): Upon derivatization, the resulting compound is a combination of the analyte with 6'-AO-TAMRA ($C_{32}H_{37}N_5O_6$), minus one H_2O unit. This increases the mass of the analyte by 569.2638 from the addition of $C_{32}H_{35}N_5O_5$.