



# A study of 9 common breath VOCs in 504 healthy subjects using PTR-TOF-MS

Zhunyan Jia<sup>1,2</sup> · Wei Qiang Ong<sup>1</sup> · Fuchang Zhang<sup>1</sup> · Fang Du<sup>1</sup> · Velmurugan Thavasi<sup>2</sup> · Venkatesan Thirumalai<sup>2</sup>

Received: 10 May 2024 / Accepted: 6 June 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

## Abstract

**Introduction** This study employs Proton-Transfer-Reaction Mass Spectrometry (PTR-MS) to analyze exhaled breath profiles of 504 healthy adults, focusing on nine common volatile organic compounds (VOCs): acetone, acetaldehyde, acetonitrile, ethanol, isoprene, methanol, propanol, phenol, and toluene. PTR-MS offers real-time VOC measurement, crucial for understanding breath biomarkers and their applications in health assessment.

**Objectives** The study aims to investigate how demographic factors—gender, age, and smoking history—affect VOC concentrations in exhaled breath. The objective is to enhance our understanding of breath biomarkers and their potential for health monitoring and clinical diagnosis.

**Methods** Exhaled breath samples were collected using PTR-MS, measuring concentrations of nine VOCs. The data were analyzed to discern distribution patterns across demographic groups.

**Results** Males showed higher average VOC levels for certain compounds. Propanol and methanol concentrations significantly increased with age. Smoking history influenced VOC levels, with differences among non-smokers, current smokers, and ex-smokers.

**Conclusion** This research provides valuable insights into demographic influences on exhaled VOC profiles, emphasizing the potential of breath analysis for health assessment. PTR-MS's real-time measurement capabilities are crucial for capturing dynamic VOC changes, offering advantages over conventional methods. These findings lay a foundation for advancements in non-invasive disease detection, highlighting the importance of considering demographics in breath biomarker research.

**Keywords** Volatile organic compounds · Breath analysis · PTR-MS · Biomarker

## 1 Introduction

Exhaled breath, an intricate amalgamation of volatile organic compounds (VOCs), has emerged as a promising source of biomarkers for non-invasive clinical diagnostics, therapeutic monitoring, and overall health assessment. Comprising nearly a thousand VOCs, including acetone, acetaldehyde, isoprene, ethanol, methanol, and various alcohols, the composition of exhaled breath reflects metabolic processes within the human body (Fenske & Paulson, 1999; Davies

et al., 2014). Over the years, extensive research has sought to identify and quantify specific VOCs in breath, aiming to leverage their diagnostic potential (Bobak et al., 2021; Li et al., 2021; Giovannini et al., 2021; Saidi et al., 2020; Maiti et al., 2020; Miller-Atkins et al., 2020; Li et al., 2020; Raspagliesi et al., 2020; Azim et al., 2019; Lawal et al., 2017; Jia et al., 2019), therapeutic monitoring capabilities (Morey et al., 2013; López-Lorente et al., 2021; Singh et al., 2021), and contributions to general wellness (Capone et al., 2018; Hamblin & Almirall, 2017; Phillips et al., 2014; Anderson, 2015). Acetone, for instance, serves as a biomarker for lipids, metabolism (Kosmider et al., 2020; Jones, 1995). Technological advancements, particularly proton-transfer reaction mass spectrometry, have played a pivotal role in enhancing the precision and sensitivity of VOC detection. This study builds upon this foundation, employing PTR-MS to delve into the breath profiles of 504 healthy adults,

✉ Venkatesan Thirumalai  
venky@ou.edu

Zhunyan Jia  
jiazhunyan@breathonix.com

<sup>1</sup> Breathonix Pte Ltd, Singapore, Singapore

<sup>2</sup> University of Oklahoma, Norman, OK, USA

systematically exploring the concentrations of nine selected common VOCs.

This study employs Proton-Transfer Reaction Mass Spectrometry (PTR-MS) as a powerful tool for the real-time measurement of volatile organic compounds (VOCs) in exhaled breath. The real-time capabilities of PTR-MS offer a significant advantage over conventional methods such as Gas Chromatography-Mass Spectrometry (GC-MS) (Herbig et al., 2009). The immediate analysis of breath samples mitigates potential issues associated with sample collection and sample transportation, reducing the risk of sample contamination and sample degradation. This methodological choice not only enhances the temporal resolution of the analysis but also contributes to the precision and efficiency of breath profiling.

The current study aims to quantify VOCs in a diverse population, unravel the intricate interplay of demographic factors such as gender, age, and smoking history on VOC concentrations, and contextualize these findings within existing literature. The overarching goal is to advance our understanding of the potential applications of breath analysis in clinical diagnosis and health monitoring. The study's comprehensive approach, encompassing a broad array of demographic factors, contributes valuable insights to the expanding body of knowledge in breath biomarkers.

## 2 Method

### 2.1 Study population

504 healthy adults, age ranging from 18 to 91 years old, were recruited in clinical trial in Nadd Al Hamar Health Center, Dubai from February to March 2021. All the participants were required to self-report through a questionnaire which included gender, age, race, smoking status, existing diseases and time for last meal. These volunteers did not undergo an overnight fasting regime and hence, the time since their last meal can range from 1 to 24 h. Each participant was instructed to breathe in and exhale normally into the PTR-TOF-MS through a disposable mouthpiece connected to the BET sampler. Three consecutive exhalations were sampled and measured in real time for each individual. The BET sampler administers a single exhalation through a customized tube, ensuring that the end-tidal fraction of the breath-gas sample is buffered (Herbig et al., 2008). Table 1 shows a demographic breakdown of the recruited adults.

### 2.2 Ethical statement

The trial was approved by the Dubai Scientific Research Ethics Committee (DSREC) of Dubai Health Authority (Ref:

**Table 1** Demographic summary of the 504 recruited adults

S/N	Classification	No. of adults
1	Age	
	< 20 years old	30
	20–60 years old	435
	> 61 years old	39
2	Gender	
	Male	285
3	< Female	219
	Smoking history	
	Non-smoker	376
	Current smoker	115
	Ex-smoker	13

**Table 2** Technical specifications of PTR-MS

Mass resolution	> 1500 m/Δm (FWHM) for m/z > 79
Response time	< 100 ms
Sensitivity	> 200 cps/ ppbv for m/z 181
Detection limits	< 10 pptv for m/z 181 (averaged over 1 min)
Linearity range	10 pptv–1 ppmv
Adjustable flow	50–800 sccm

DSREC-01/2021\_07) and all enrolled subjects gave signed informed consent prior to the study.

### 2.3 PTR-MS measurement & analysis

Breath samples were measured by a PTR-MS TOF1000, Ionicon Analytik GmbH, Innsbruck, Austria. PTR-TOF-MS consists of an ionization section and a detection section. Within the ionization process, protonated water ions ( $\text{H}_3\text{O}^+$ ) were generated through a hollow cathode discharge in the ion source. These  $\text{H}_3\text{O}^+$  ions were subsequently introduced into the drift tube by an electric drift field, facilitating the chemical ionization of volatile organic compounds (VOCs) in breath samples via proton-transfer reactions (PTR). Only VOCs with a higher proton affinity (PA) value than that of  $\text{H}_2\text{O}$  molecules underwent ionization by  $\text{H}_3\text{O}^+$  and proceeded to the detection section. The ionized VOCs were then directed by an electric field towards the time-of-flight mass spectrometer (TOF-MS), where they were differentiated and detected based on their mass-to-charge ratio (m/z) (Table 2).

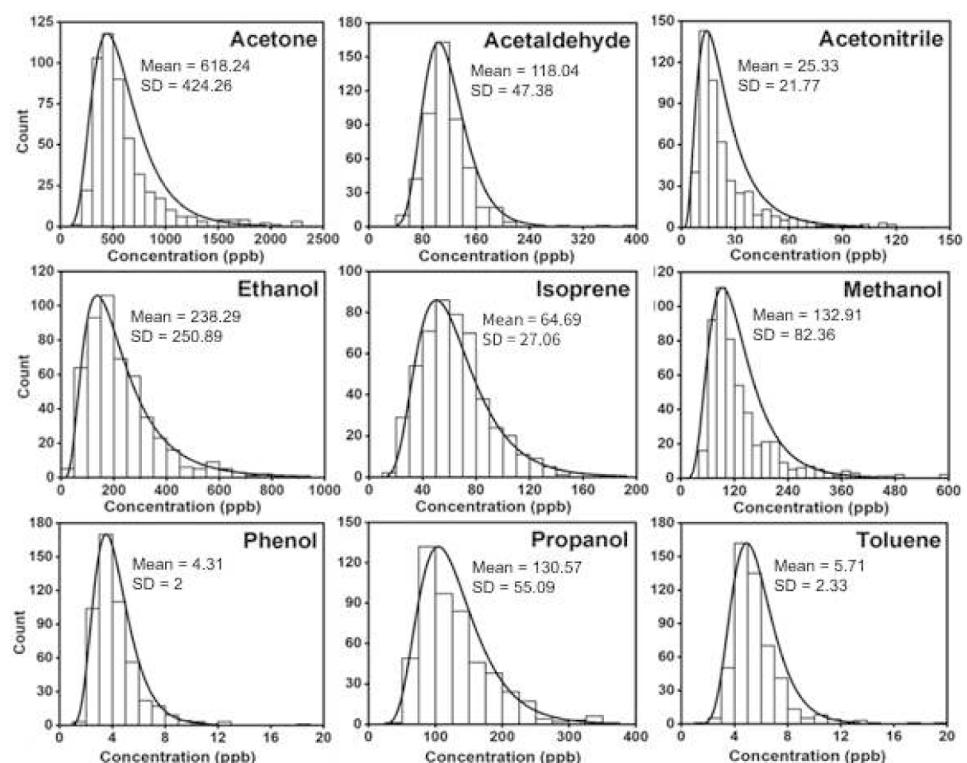
Key operational parameters of the PTR-MS instrument included a drift tube voltage of 600V, a temperature of 60 °C, a drift tube pressure of 2.3 mbar, and an E/N ratio of 139 Td. Additionally, the sampling line and buffer tube were maintained at 70 °C. Preceding each breath sample acquisition, a 5-second background concentration measurement was performed to establish a reference point for subsequent

analysis. Participants were then instructed to breathe naturally through a mouthpiece, allowing sufficient time for VOC levels to equilibrate with the background concentration before advancing to the subsequent sample. This procedure was iterated three times per participant, and the resulting dataset was averaged to generate reliable and representative VOC concentrations suitable for subsequent analysis.

Mass calibration was conducted using masses 21.0226 (H<sub>3</sub>O<sup>+</sup> isotope), 60.049 (acetone isotope), and 203.943 (diiodobenzene fragment), once every 60 s. VOC calibration for the Ionicon PTR-MS system was performed monthly using a standard gas cylinder.

The PTR-MS will calculate the concentration of all VOCs and save it in .h5 files (Herbig et al., 2009). This raw concentration data was processed by software, Viewer software 4.2.0, Ionicon Analytik GmbH, Innsbruck, Austria, for mass calibration and peak data calculation. A trace analyzer, Viewer software 4.2.0, was subsequently employed to identify background air and exhaled breath, with data points from the end-tidal phase being selected and averaged for the three exhalations from each subject. VOCs exhibiting concentrations lower in breath than in the background were excluded from subsequent data analysis. An Excel file containing the list of VOCs and their concentrations (in parts per billion, ppb) was generated for each sample. Finally, the VOC concentration data was analysed by R.

**Fig. 1** Distributions of the concentrations in parts-per-billion (ppb) of the 9 common VOCs of exhaled breath examined in this study. The geometric mean (GM) and geometric standard deviation (GSD) of each VOC are shown in the respective plots



### 3 Results and discussion

#### 3.1 Concentration distribution of the exhaled VOCs

The examination of exhaled breath from 504 healthy adults using proton-transfer-reaction mass spectrometry (PTR-MS) revealed a concentration distribution pattern for the nine common volatile organic compounds (VOCs), the concentrations of these VOCs—acetone, acetaldehyde, acetonitrile, ethanol, isoprene, methanol, propanol, phenol, and toluene—closely, as illustrated in Fig. 1.

Acetone, a byproduct of lipid metabolism and a recognized biomarker for fat loss (Anderson, 2015), exhibited an overall concentration range in adults from 176 to 4321 ppb, with median and geometric mean concentrations at 512 ppb and 545 ppb, respectively.

Breath acetaldehyde, serving as a tool to investigate exposures to smoking (Tardif, 2006), ethanol, or acetaldehyde, displayed a concentration range of 40 to 396 ppb. The median and geometric mean for acetaldehyde were consistent at 113 ppb, with a notable outlier at 843 ppb—an individual identified as a heavy current smoker.

Adults were exposed to acetonitrile through breathing, as it is present in cigarette smoke and automobile exhaust. The overall concentration range of acetonitrile in this study's adults was 5.99 to 167 ppb, with a median and geometric

mean at 17.6 ppb and 20.4 ppb, respectively (Wong et al., 1992).

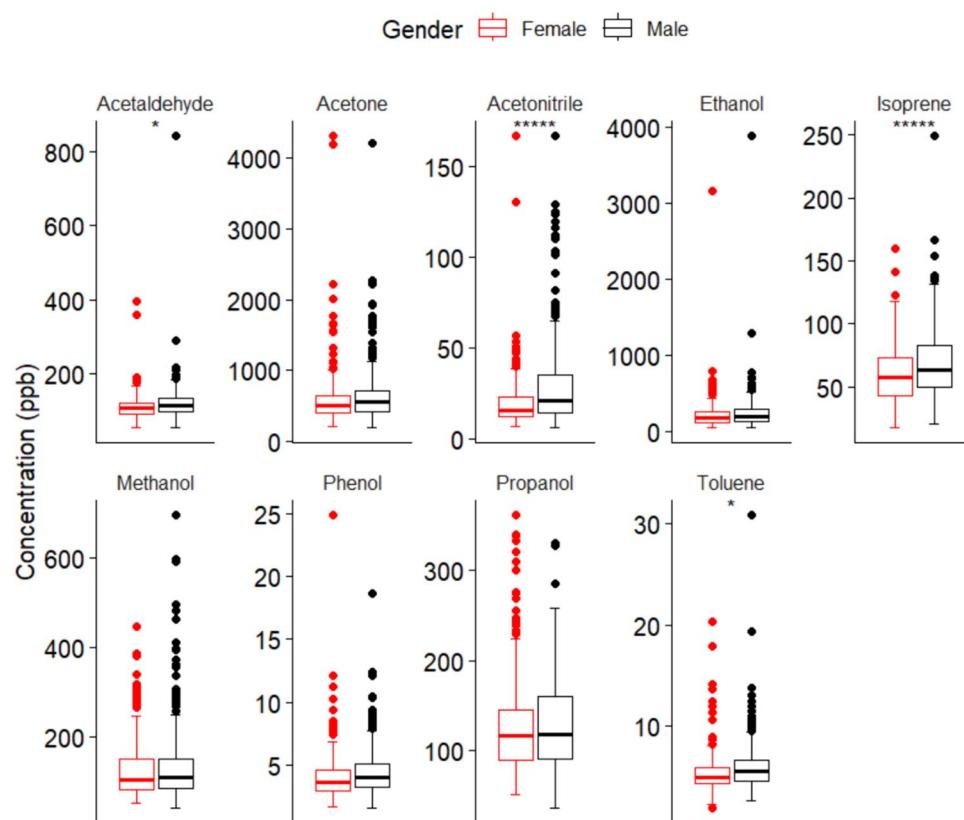
Among the alcohols-ethanol, methanol, and propanol-major components found in breath, ethanol exhibited the highest abundance (Issitt et al., 2022). The overall concentration range of ethanol spanned from 44 to 3891 ppb, with a geometric mean and median both at 194 and 191 ppb, respectively. In the case of methanol, the concentration range was from 41 to 696 ppb, approximately 5.6 times lower than ethanol, with a geometric mean and median at 117 and 107 ppb, respectively. Propanol, with a concentration range of 36 to 362 ppb, displayed lower concentrations, and the geometric mean and median were 121 and 116 ppb, respectively.

Isoprene, a major VOC in breath and a versatile biomarker for health monitoring (Fenske & Paulson, 1999),

**Table 3** Average concentration of VOCs that were significantly different between male and female on ANOVA test ( $\alpha = 0.05$ )

VOC	Means (PPB)		SD (PPB)		ANOVA test's p value
	Male	Female	Male	Female	
Acetaldehyde	122	113	53	37	0.022
Acetonitrile	30	20	25	16	< 0.001
Isoprene	69	59	29	23	< 0.001
Toluene	5.9	5.4	2.5	2.1	0.024

**Fig. 2** Box/whisker plots showing the gender differences of VOCs in exhaled breath. For all the VOCs, men had a higher median VOC concentration as compared to women. \*Significant differences on ANOVA test ( $\alpha = 0.05$ )



exhibited an overall concentration range from 17 to 249 ppb, with a median and geometric mean at 61 and 59 ppb, respectively.

Phenol and toluene, aromatic compounds with uncertain endogenous origins within the body, exhibited breath concentrations ranging from 1.6 to 46 ppb for phenol and 1.9 to 40.9 ppb for toluene. The median and geometric mean concentrations were 3.8 and 4.0 ppb for phenol, and 5.2 and 5.4 ppb for toluene, respectively. This distribution pattern observed across the analyzed volatile organic compounds (VOCs) lays a foundational understanding of their prevalence in exhaled breath, thereby contributing valuable insights to the ongoing investigation of breath biomarkers.

### 3.2 Effects of gender on exhaled VOCs

The investigation into the influence of gender on exhaled volatile organic compounds (VOCs) revealed noteworthy disparities in the average VOC concentrations between men and women. As depicted in Fig. 2 and substantiated by analysis of variance (ANOVA), men consistently exhibited higher median VOC concentrations compared to women across all examined VOCs. ANOVA results indicated significant gender-based differences for exhaled acetaldehyde, acetonitrile, isoprene, and toluene, even with variations in sample sizes. Table 3 provides a summary of the mean concentrations and

ANOVA test p-values for these VOCs. Notably, the breath isoprene concentration in men (mean 69 ppb) was significantly higher than in women (mean 59 ppb), with a considerable difference ( $p < 0.001$ ). This finding aligns with previous research demonstrating gender-related effects on isoprene levels (Lechner et al., 2006). Conversely, ANOVA analysis revealed no significant gender-based differences for breath acetone, ethanol, methanol, propanol, and phenol. While men generally displayed higher mean concentrations for acetone, a departure from previous studies (Turner et al., 2006), the absence of a significant difference in our study could be attributed to the non-restriction of participants' overnight fasting. The box/whisker plots in Fig. 2 illustrate the gender differences in VOC concentrations, emphasizing the importance of nuanced considerations in demographic factors during breath analysis. These findings contribute to a more refined understanding of gender-related variations in breath biomarkers, offering insights essential for future research in clinical applications and health monitoring.

### 3.3 Effects of age on exhaled VOCs

An exploration into the impact of age on exhaled volatile organic compounds (VOCs) revealed discernible patterns in VOC concentrations across different age groups. As delineated in Table 4 and Fig. 3, the study categorized participants into three age groups: those below 20 years old, individuals between 20 and 60 years old, and those surpassing 60 years old. For acetone, acetaldehyde, acetonitrile, isoprene, and toluene, the 20 to 60 years old age group consistently exhibited the highest mean concentrations compared to the younger and older groups. While the differences were not always statistically significant, a notable trend emerged.

Propanol emerged as a unique VOC, displaying a significant increase in concentration with age ( $p = 0.0013$ ), as illustrated in Fig. 3a. The average propanol concentrations for participants less than 20 years old, between 20 and 60 years old, and over 60 years old were 110, 129, and 157 ppb, respectively. Regression modelling supported

**Table 4** Mean & Coefficient of Variance (COV) of VOCs concentration of different age groups

VOC	Mean/ppb (COV/%)		
	< 20 years old	20–60 years old	> 60 years old
Acetone	561 (44)	624 (72)	594 (37)
Acetaldehyde	118 (22)	118 (42)	116 (30)
Acetonitrile	18 (51)	26(87)	21 (64)
Ethanol	253 (52)	237(112)	244 (57)
Isoprene	58 (34)	66(42)	57 (38)
Methanol	107 (45)	134 (63)	145 (52)
Propanol*	110 (35)	130 (42)	157 (40)
Phenol	4.1 (25)	4.3 (48)	4.4 (46)
Toluene	4.9 (20)	5.8 (42)	5.6 (35)

COV is calculated by dividng the sample standard deviation by sample mean

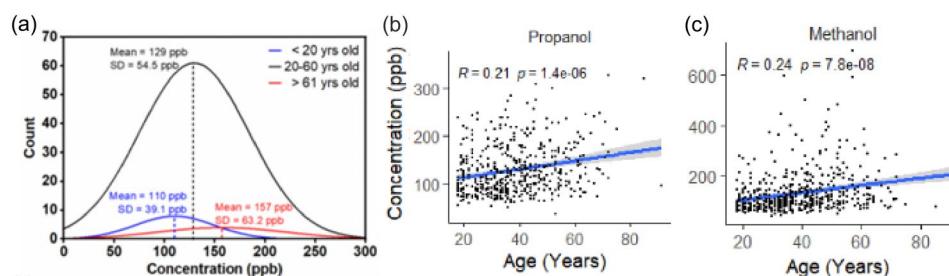
\*Significant difference on ANOVA test ( $\alpha = 0.05$ )

this observation, revealing a positive correlation between propanol concentration and age ( $p < 0.001$ ), with a slope of  $0.8677 (\pm 0.1774)$  ppb/year (Fig. 3b). Similarly, a positive correlation was observed for methanol, indicating an increase in concentration with age ( $p < 0.001$ ). In the dataset, the methanol concentration was noted to increase by  $1.4383 (\pm 0.2638)$  ppb for every year of age (Fig. 3c).

In contrast, acetaldehyde concentrations exhibited relative stability across various age brackets, indicating that age may not exert a substantial impact on this specific volatile organic compound (VOC). These results offer fresh perspectives on age-related shifts in breath composition, highlighting the importance of propanol and methanol, and offering valuable insights for prospective studies in clinical diagnostics and health surveillance.

### 3.4 Effects of smoking on exhaled VOCs

In the dataset, significant differences were observed in the concentrations of breath acetone, isoprene, and toluene

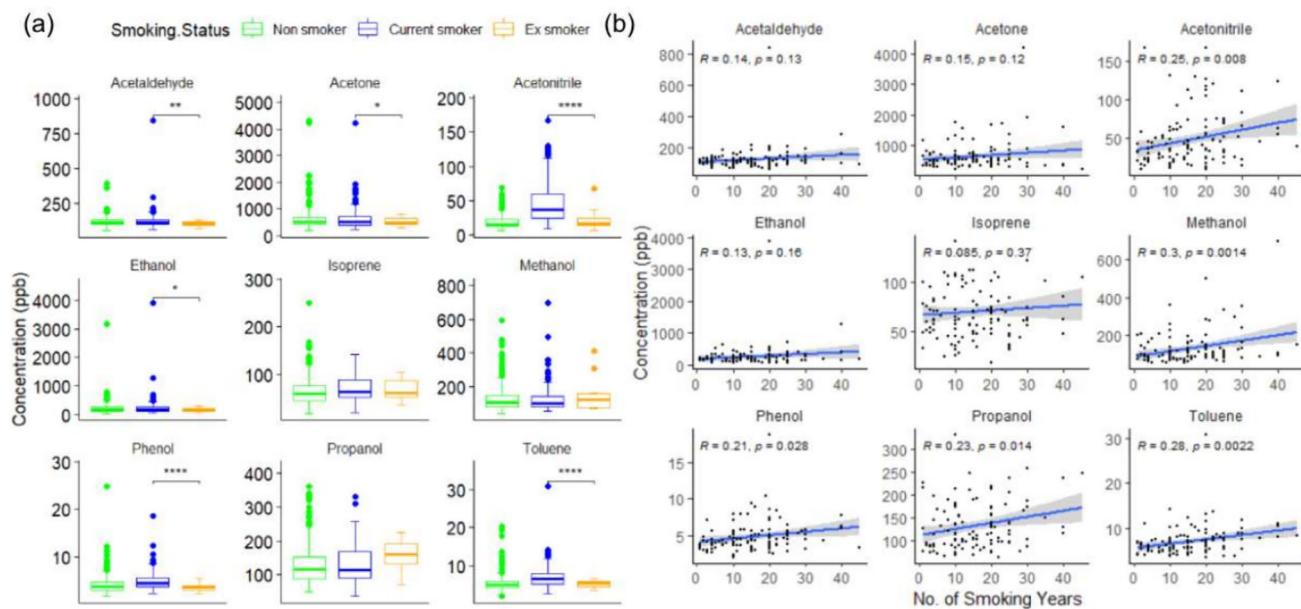


**Fig. 3** a Histograms showing the significant difference in distribution of propanol concentration of the 3 different age groups, i.e. less than 20 years old (blue line), 20 to 60 years old (black line) and

greater than 60 years old (red line). The mean and standard deviation (SD) of each group are shown in the respective colours. Correlation of (b) propanol and (c) methanol concentration with age (years)

between smoking and non-smoking groups. Toluene, a recognized biomarker of smoking (Capone et al., 2018), exhibited noteworthy distinctions among non-smokers, current smokers, and ex-smokers. The average breath acetone concentrations for non-smokers, current smokers, and ex-smokers were determined to be 611, 654, and 516 ppb, respectively. While the ANOVA test did not reveal significant differences among the groups (Fig. 4a), our findings align with a previously reported study indicating higher acetone concentrations in smokers compared to non-smokers (Capone et al., 2018). Notably, ex-smokers exhibited the lowest acetone concentrations, a trend consistent with acetaldehyde, ethanol, and phenol. Similarly, the average breath isoprene concentration of current smokers (mean 71 ppb) exceeded that of non-smokers (mean 63 ppb) (Fig. 4a). This result is in accordance with previous reports suggesting that the breath isoprene levels of smokers can be up to 8 times higher than those of non-smokers (Capone et al., 2018; Kushch et al., 2008). Following smoking cessation, the breath isoprene concentration decreased, yet ex-smokers maintained a level (mean 66 ppb) higher than that of non-smokers, and ANOVA testing indicated significant differences among the three groups ( $p = 0.028$ ). The average toluene concentrations for non-smokers, current smokers, and ex-smokers were 5.3, 7.0, and 5.3 ppb, respectively (Fig. 4a). Given the exposure of smokers to small amounts of toluene in cigarette smoke, the significantly higher breath toluene concentration in current smokers compared to non-smokers and ex-smokers was not surprising ( $p < 0.001$ ). This finding corroborates previous studies (Capone et al., 2018; Kushch

et al., 2008). A positive correlation was observed between breath toluene concentration and years of smoking (Fig. 4b). For each year of smoking, the breath toluene concentration was determined to increase by 0.0861 ( $\pm 0.0216$ ) ppb, translating to a 1.1-fold increase for every 10 years of smoking. Phenol, another aromatic compound distinct from toluene, is not widely recognized as a smoking biomarker and has not found a significant difference between smokers and non-smokers Marco and Grimalt (2015). Consequently, the significant difference observed in breath phenol concentrations among non-smokers, current smokers, and ex-smokers ( $p < 0.001$ ) was intriguing (Fig. 4a). Current smokers exhibited an average breath phenol concentration of 4.9 ppb, surpassing that of non-smokers (mean 4.2 ppb). In contrast, ex-smokers displayed a markedly lower average phenol concentration (mean 3.4 ppb) compared to non-smokers. A positive correlation was identified between breath phenol concentration and the number of smoking years ( $p = 0.032$ ) (Fig. 4b). An increase of 0.0434 ( $\pm 0.0200$ ) ppb in phenol concentration was observed for each year of smoking, equivalent to a 1.1-fold increase for every 10 years of smoking. Both acetaldehyde and acetonitrile are components of cigarette smoke (Seeman et al., 2002), and these VOCs exhibited similar trends (Fig. 4a). Current smokers demonstrated a significantly higher acetaldehyde concentration (mean 127 ppb) compared to non-smokers (mean 116 ppb;  $p = 0.036$ ). Ex-smokers exhibited a mean acetaldehyde concentration of 103 ppb, lower than that of non-smokers, although the difference was not statistically significant. Likewise, the breath acetonitrile concentrations among the three smoking



**Fig. 4** **a** Boxwhisker plots of selected VOCs showing the effects of smoking on VOCs in exhaled breath. \*Significant differences on ANOVA test ( $\alpha = 0.05$ )). **b** Correlation of exhaled breath VOCs concentration of current smoker with number of years of smoking

**Table 5** Differences between the ‘smoking history’ groups to show the effect of smoking on the average breath alcohol concentration

Differences between groups	Current smoker non smoker	Ex-smoker current smoker
Concentration/ppb	Ethanol	27
	Methanol	1
	Propanol	4
		88
		14
		27

history groups were significantly different ( $p < 0.001$ ), with non-smokers, current smokers, and ex-smokers having mean concentrations of 19, 47, and 22 ppb, respectively. Current smokers had a breath acetonitrile level approximately 2.5 times higher than that of non-smokers, aligning with previous studies reporting elevated acetonitrile concentrations in smokers (Kushch et al., 2008). Following smoking cessation, ex-smokers exhibited a breath acetonitrile level about 1.2 times that of non-smokers. Previous research has indicated that within approximately a week of smoking cessation, the breath acetonitrile concentration drops to levels comparable to those of non-smokers (Jordan et al., 1995), consistent with our findings. Interestingly, a positive correlation was observed between mean breath acetaldehyde ( $p = 0.020$ ) or acetonitrile ( $p = 0.008$ ) concentrations in current smokers and the number of years they had been smoking (Fig. 4b). The regression line slope for mean breath acetaldehyde concentration against the number of years smoking was found to be  $0.8259 (\pm 0.3497)$  ppb per year of smoking, corresponding to a 1.1-fold increase in breath acetaldehyde level for every 10 years of smoking. Similarly, for acetonitrile, there was an increase of  $0.8843 (\pm 0.3274)$  ppb in breath concentration for every year of smoking, resulting in a 1.2-fold increase in acetonitrile concentration over 10 years of smoking.

The three alcohols under investigation-methanol, propanol, and ethanol-do not serve as established biomarkers of smoking. Consequently, as anticipated, no significant differences were observed among the three smoking groups. Generally, current smokers exhibited a higher average breath alcohol concentration compared to non-smokers. However, it is noteworthy to highlight the divergent trends influenced by smoking on breath alcohol concentration. Smoking’s impact on the breath of methanol and propanol mirrored that of ethanol in a similar manner (Table 5). Specifically, marginal differences in average concentrations were noted between non-smokers and current smokers for methanol and propanol (1 and 4 ppb, respectively). In contrast, a substantial difference of 27 ppb was observed between these two groups for breath ethanol. Upon comparing ex-smokers with current smokers, a considerable increase in average breath methanol and propanol concentrations in ex-smokers, differing by 4 and 27 ppb, respectively, was noted. However, in the case

of ethanol, ex-smokers exhibited a significantly lower average breath concentration than both current and non-smokers, with a notable difference of 88 ppb. These variations in ethanol trends could potentially be attributed to its involvement in more intricate metabolic processes within the body.

In tandem with these advancements, it is crucial to recognize the limitations of Proton Transfer Reaction Mass Spectrometry (PTR-MS). Despite its potential, PTR-MS exhibits several drawbacks. Its precision is comparatively modest, making it challenging to differentiate between molecules with similar molecular weights. Furthermore, PTR-MS can only identify molecular weights, lacking specificity in naming individual molecules. The accuracy of concentration calibration varies across different intervals, and does not match the high precision observed in Gas Chromatography-Mass Spectrometry (GC-MS). Additionally, batch effects, particularly those related to device variations across different locations, pose inherent challenges.

## 4 Conclusions

In summary, the analysis of exhaled breath profiles from a cohort comprising 504 healthy adults, conducted using PTR-MS in this study, not only confirms existing research but also introduces fresh perspectives on how age, gender, and smoking history impact volatile organic compounds (VOCs). Our findings shed light on age-related variations, notably the significant rise in propanol concentrations associated with increasing age, thereby enhancing our understanding of the physiological determinants influencing breath composition. Gender-specific differences in VOC concentrations, particularly evident for acetaldehyde, acetonitrile, isoprene, and toluene, highlight the importance of considering demographic factors with subtlety in breath analysis. The inclusion of ex-smokers in our study extends insights into the long-term effects of smoking cessation on breath VOCs, a facet often overlooked in prior studies. These initial findings underscore the potential of breath analysis as a non-invasive method for health monitoring and clinical diagnosis, setting the stage for future advancements that could transform personalized healthcare through the nuanced exploration of breath biomarkers.

**Acknowledgements** V. Thirumalai and V. Thavasi would like to express their sincere gratitude to the National Science Foundation for their grant (2201561, ‘PIPP Phase I’), as well as to the Vice President of Research and Partnerships at OU and OUHSC for their generous support. They also extend thanks to OU’s Data Institute for Societal Challenges (DISC) for its significant support.

**Author contributions** Conceptualization, Zhunan Jia and Venkatesan Thirumalai; methodology, Zhunan Jia and Wei Qiang Ong; software, Fuchang Zhang; formal analysis, Zhunan Jia and Fuchang Zhang; writing-original draft preparation, Wei Qiang Ong and Zhunan Jia;

writing-review and editing, Wei Qiang Ong, Zhunan Jia, Velmurugan Thavasi and Venkatesan Thirumalai; visualization, Fuchang Zhang; project administration, Fang Du; funding acquisition, Zhunan Jia and Fang Du. All authors have read and agreed to the published version of the manuscript.

**Funding** This research received no external funding.

**Data availability** This study is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench, <https://www.metabolomicsworkbench.org> where it has been assigned Study ID ST003200. The data can be accessed directly via its Project <http://dx.doi.org/10.21228/M8014R>.

**Code availability** The software developed in this study is available via [inset URL].

## Declarations

**Competing interests** The authors declare no competing interests.

**Informed consent** This study was approved by the Nadd Al Hamar Health center institutional review board. All participants gave informed consent before taking part in the breath analysis. Patients signed informed consent regarding publishing their data and photographs.

## References

Anderson, J. C. (2015). Measuring breath acetone for monitoring fat loss: Review. *Obesity*, 23(12), 2327–2334. <https://doi.org/10.1002/oby.21242>

Azim, A., Barber, C., Dennison, P., Riley, J., & Howarth, P. (2019). Exhaled volatile organic compounds in adult asthma: A systematic review. *European Respiratory Journal*, 54(3), 1900056. <https://doi.org/10.1183/13993003.00056-2019>

Bobak, C. A., Kang, L., Workman, L., Bateman, L., Khan, M. S., Prins, M., May, L., Franchina, F. A., Baard, C., Nicol, M. P., Zar, H. J., & Hill, J. E. (2021). Breath can discriminate tuberculosis from other lower respiratory illness in children. *Scientific Reports*, 11(1), 2704. <https://doi.org/10.1038/s41598-021-80970-w>

Capone, S., Tufariello, M., Forleo, A., Longo, V., Giampetrucci, L., Radogna, A. V., Casino, F., & Siciliano, P. (2018). Chromatographic analysis of VOC patterns in exhaled breath from smokers and nonsmokers. *Biomedical Chromatography*, 32(4), 4132. <https://doi.org/10.1002/bmc.4132>

Davies, S. J., Španěl, P., & Smith, D. (2014). Breath analysis of ammonia, volatile organic compounds and deuterated water vapor in chronic kidney disease and during dialysis. *Bioanalysis*, 6(6), 843–857. <https://doi.org/10.4155/bio.14.26>

Fenske, J. D., & Paulson, S. E. (1999). Human breath emissions of VOCs. *Journal of the Air & Waste Management Association*, 49(5), 594–598. <https://doi.org/10.1080/10473289.1999.10463831>

Giovannini, G., Haick, H., & Garoli, D. (2021). Detecting COVID-19 from breath: A game changer for a big challenge. *ACS Sensors*, 6(4), 1408–1417. <https://doi.org/10.1021/acssensors.1c00312>

Hamblin, D., & Almirall, J. (2017). Analysis of exhaled breath from cigarette smokers using CMV-GC/MS. *Forensic Chemistry*, 4, 67–74. <https://doi.org/10.1016/j.forc.2017.02.008>

Herbig, J., Müller, M., Schallhart, S., Titzmann, T., Graus, M., & Hansel, A. (2009). On-line breath analysis with PTR-TOF. *Journal of Breath Research*, 3(2), 027004.

Herbig, J., Titzmann, T., Beauchamp, J., Kohl, I., & Hansel, A. (2008). Buffered end-tidal (bet) sampling—A novel method for real-time breath-gas analysis. *Journal of Breath Research*, 2(3), 037008.

Issitt, T., Wiggins, L., Veysey, M., Sweeney, S. T., Brackenbury, W. J., & Redeker, K. (2022). Volatile compounds in human breath: Critical review and meta-analysis. *Journal of Breath Research*, 16(2), 024001. <https://doi.org/10.1088/1752-7163/ac5230>

Jia, Z., Patra, A., Kutty, V., & Venkatesan, T. (2019). Critical review of volatile organic compound analysis in breath and in vitro cell culture for detection of lung cancer. *Metabolites*, 9(3), 52. <https://doi.org/10.3390/metabo9030052>

Jones, A. W. (1995). Measuring and reporting the concentration of acetaldehyde in human breath. *Alcohol and Alcoholism*, 30(3), 271–285. <https://doi.org/10.1093/oxfordjournals.alcalc.a045731>

Jordan, A., Hansel, A., Holzinger, R., & Lindinger, W. (1995). Acetonitrile and benzene in the breath of smokers and non-smokers investigated by proton transfer reaction mass spectrometry (PTR-MS). *International Journal of Mass Spectrometry and Ion Processes*, 148(1–2), 1–3. [https://doi.org/10.1016/0168-1176\(95\)04236-E](https://doi.org/10.1016/0168-1176(95)04236-E)

Kosmider, L., Cox, S., Zaciura, M., Kurek, J., Goniewicz, M. L., McRobbie, H., Kimber, C., & Dawkins, L. (2020). Daily exposure to formaldehyde and acetaldehyde and potential health risk associated with use of high and low nicotine e-liquid concentrations. *Scientific Reports*, 10(1), 6546. <https://doi.org/10.1038/s41598-020-63292-1>

Kushch, I., Schwarz, K., Schwentner, L., Baumann, B., Dzien, A., Schmid, A., Unterkofler, K., Gastl, G., Španěl, P., Smith, D., & Amann, A. (2008). Compounds enhanced in a mass spectrometric profile of smokers' exhaled breath versus non-smokers as determined in a pilot study using PTR-MS. *Journal of Breath Research*, 2(2), 026002. <https://doi.org/10.1088/1752-7155/2/2/026002>

Lawal, O., Ahmed, W. M., Nijssen, T. M. E., Goodacre, R., & Fowler, S. J. (2017). Exhaled breath analysis: A review of 'breath-taking' methods for off-line analysis. *Metabolomics*, 13(10), 110. <https://doi.org/10.1007/s11306-017-1241-8>

Lechner, M., Moser, B., Niederseer, D., Karlseder, A., Holzknecht, B., Fuchs, M., Colvin, S., Tilg, H., & Rieder, J. (2006). Gender and age specific differences in exhaled isoprene levels. *Respiratory Physiology & Neurobiology*, 154(3), 478–483. <https://doi.org/10.1016/j.resp.2006.01.007>

Li, J., Guan, X., Fan, Z., Ching, L.-M., Li, Y., Wang, X., Cao, W.-M., & Liu, D.-X. (2020). Non-invasive biomarkers for early detection of breast cancer. *Cancers*, 12(10), 2767. <https://doi.org/10.3390/cancers12102767>

Li, X., Li, J., Ge, Q., Du, Y., Li, G., Li, W., Zhang, T., Tan, L., Zhang, R., Yuan, X., Zhang, H., Zhang, C., Liu, W., Ding, W., Sun, L., Chen, K., Wang, Z., Shen, N., & Lu, J. (2021). Detecting SARS-CoV-2 in the breath of COVID-19 patients. *Frontiers in Medicine*, 8, 604392. <https://doi.org/10.3389/fmed.2021.604392>

López-Lorente, C. I., Awchi, M., Sinues, P., & García-Gómez, D. (2021). Real-time pharmacokinetics via online analysis of exhaled breath. *Journal of Pharmaceutical and Biomedical Analysis*, 205, 114311. <https://doi.org/10.1016/j.jpba.2021.114311>

Maiti, K. S., Roy, S., Lampe, R., & Apolonski, A. (2020). Breath indeed carries significant information about a disease: Potential biomarkers of cerebral palsy. *Journal of Biophotonics*, 13(11), 202000125. <https://doi.org/10.1002/jbio.202000125>

Marco, E., & Grimalt, J. O. (2015). A rapid method for the chromatographic analysis of volatile organic compounds in exhaled breath of tobacco cigarette and electronic cigarette smokers. *Journal of Chromatography A*, 1410, 51–59. <https://doi.org/10.1016/j.chroma.2015.07.094>

Miller-Atkins, G., Acevedo-Moreno, L.-A., Grove, D., Dweik, R. A., Tonelli, A. R., Brown, J. M., Allende, D. S., Aucejo, F., & Rotroff, D. M. (2020). Breath metabolomics provides an accurate and non-invasive approach for screening cirrhosis, primary, and secondary

liver tumors. *Hepatology Communications*, 4(7), 1041–1055. <https://doi.org/10.1002/hep4.1499>

Morey, T. E., Booth, M., Wasdo, S., Wishin, J., Quinn, B., Gonzalez, D., Derendorf, H., McGorray, S. P., Simoni, J., Melker, R. J., & Dennis, D. M. (2013). Oral adherence monitoring using a breath test to supplement highly active antiretroviral therapy. *AIDS and Behavior*, 17(1), 298–306. <https://doi.org/10.1007/s10461-012-0318-7>

Phillips, M., Beatty, J. D., Cataneo, R. N., Huston, J., Kaplan, P. D., Lalisang, R. I., Lambin, P., Lobbes, M. B. I., Mundada, M., Pappas, N., & Patel, U. (2014). Rapid point-of-care breath test for biomarkers of breast cancer and abnormal mammograms. *PLoS ONE*, 9(3), 90226. <https://doi.org/10.1371/journal.pone.0090226>

Raspagliesi, F., Bogni, G., Benedetti, S., Grassi, S., Ferla, S., & Buratti, S. (2020). Detection of ovarian cancer through exhaled breath by electronic nose: A prospective study. *Cancers*, 12(9), 2408. <https://doi.org/10.3390/cancers12092408>

Saidi, T., Moufid, M., De Jesus Beleño-Saenz, K., Welearegay, T. G., El Bari, N., Lisset Jaimes-Mogollon, A., Ionescu, R., Bourkadi, J. E., Benamor, J., El Ftouh, M., & Bouchikhi, B. (2020). Non-invasive prediction of lung cancer histological types through exhaled breath analysis by UV-irradiated electronic nose and GC/QTOF/MS. *Sensors and Actuators B: Chemical*, 311, 127932. <https://doi.org/10.1016/j.snb.2020.127932>

Seeman, J. I., Dixon, M., & Haussmann, H.-J. (2002). Acetaldehyde in mainstream tobacco smoke: Formation and occurrence in smoke and bioavailability in the smoker. *Chemical Research in Toxicology*, 15(11), 1331–1350. <https://doi.org/10.1021/tx020069f>

Singh, K. D., Osswald, M., Ziesenitz, V. C., Awchi, M., Usemann, J., Imbach, L. L., Kohler, M., García-Gómez, D., Van Den Anker, J., Frey, U., Datta, A. N., & Sinues, P. (2021). Personalised therapeutic management of epileptic patients guided by pathway-driven breath metabolomics. *Communications Medicine*, 1(1), 21. <https://doi.org/10.1038/s43856-021-00021-3>

Tardif, R. (2006). The determination of acetaldehyde in exhaled breath. In D. J. Chadwick & J. Goode (Eds.), *Novartis foundation symposia* (1st ed., Vol. 285, pp. 125–136). Wiley. <https://doi.org/10.1002/9780470511848.ch9>

Turner, C., Španěl, P., & Smith, D. (2006). A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiological Measurement*, 27(7), 637–648. <https://doi.org/10.1088/0967-3334/27/7/007>

Wong, M. K. C., Keyes Scott, B., & Peterson, C. M. (1992). Breath acetaldehyde following ethanol consumption. *Alcohol*, 9(3), 189–192. [https://doi.org/10.1016/0741-8329\(92\)90051-B](https://doi.org/10.1016/0741-8329(92)90051-B)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.