Supporting information

# Combination of exhaled breath analysis with parallel lung function and FeNO measurements in infants

Fabienne Decrue[1,2], Kapil Dev Singh[1,3], Amanda Gisler[1,2], Mo Awchi[1,3], Jiafa Zeng[1,3], Jakob Usemann[1,2], Urs Frey[1,2,3] and Pablo Sinues[1,3,*]

[1] University of Basel Children's Hospital (UKBB), Basel, 4056, Switzerland

[2] Pediatric Respiratory Medicine, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Bern, 3010, Switzerland

[3] University of Basel, Department of Biomedical Engineering, Allschwil, 4123, Switzerland

* pablo.sinues@ukbb.ch

**Table of content**

## EXPERIMENTAL SECTION

### *Comparison between real-time and offline exhaled metabolic profiles in adults*

*Study participants*

In total 13 healthy adult participants (mean age 33.8 years, range 24-43 years, 62% female) included in the study were part of a study with the aim to standardize procedures for breath analysis by SESI-HRMS. Adults provided one to two pairs of real-time and offline measurements per day with at least 4 hours of rest in between, at least one hour of fasting and abstaining from chewing gum or brushing their teeth prior to the measurement. This resulted in 190 pairs of real-time and offline measurements, from which 14 pairs had to be excluded due to technical issues (n = 8; clogged electrospray capillary leading to unstable signals) or non-adherence to study protocol (n = 6). Anthropometric data for adults is presented in ***Table S1***. The local ethics committees (Ethics Committee for Northwest and Central Switzerland) approved the study (EKBB-Nr. 360/11 and ID 2018-01324), and written informed consent was obtained.

*Breath analysis*

Real-time measurements in adults were performed following the same procedure as previously described [1] (***Figure 1***). Briefly, the breath interface (Exhalion, FIT, Spain) was used to control flow rate and volume of exhalation maneuvers (via exhaled $CO_2$), maximizing the reproducibility across individuals [2]. For each real-time measurement, six consecutive exhalations were acquired. Participants were asked to inhale through their nose and then exhale through a disposable bacterial/viral filter. Exhalion interface was coupled to a SESI ion source (SuperSESI, FIT, Spain) and a HRMS (Q-Exactive Plus, Thermo Fisher Scientific, Germany), as previously described [3,4]. The Q Exactive Plus MS was operated via Q Exactive Tune software in positive polarity full scan mode, over the mass-to-charge range of 70-1000 with a resolution of 140,000, 4 microscans, automatic gain control target $10^6$ and maximum injection time of 500 ms.

Offline breath measurements were subsequently provided within 15 min after real-time breath measurements (***Figure 1***). The offline device consisted of a one-way valve as a mouthpiece (Hudson RCI®), a Nalophan® bag (Nalophan® NA, 20 µm ± 5 µm thick, Kalle Group, Germany) of approximately 2 L volume and 700 $cm^2$ surface and a tube (Rotilabo® PTFE, 6 mm, 8 mm, length 90 mm, Carl Roth®) connected to a valve (VHK2-08F-08F, SMC Switzerland) at the end of the bag. After offline collection of breath, the samples were infused into the ion source within 10 mins after collection. The ion source featured a low-pressure mass flow controller that ensured same flow conditions through the ion source for the real-time and the offline measurements.

*Quality control*

In the morning before breath measurements, the performance of the SESI-HRMS was checked by measuring a gas standard including 8 components (Dalian Special Gases, Dalian, China; ***Table S2***) diluted to 2 ppb (dilutor Model 2010, Sabio Environmental, Round Rock, US). For all measurements, it was confirmed that the signal intensity of protonated α-Terpinene (m/z 137) was greater than $10^7$ a.u. In addition, each new measurement was compared against historic data obtained for the same gas standard and confirmed there were no significant deviations from previous data points.

*Data analysis*

Data pre-processing and further statistical analyses were performed using MATLAB (version 2020b, MathWorks Inc., USA). MS raw data was accessed via inhouse C# console apps based on Thermo Fisher Scientific's RawFileReader (version 5.0.0.38). After binning, the time traces for each mass spectra feature were extracted. The area under the curve during the exhalation windows for each of the features detected was computed and normalized to the time window.

The replicate exhalations within each real-time experiment were averaged. A similar procedure was used to calculate the signal intensities for the offline analysis. As a result, for the adults a data matrix of 352 samples (176 real-time and 176 offline) × 2,284 features was obtained. The data were 5$^{th}$-root transformed to approach normal distributions. Lin's concordance correlation coefficient [5] (Lin's CCC) was used to estimate the agreement of the offline measurements against the real-time technique in adults.

**Deployment of offline method in infants, along with lung function and fractional exhaled nitric oxide (FeNO) in infants**

*Study participants*

From the prospective and ongoing Basel-Bern Infant Lung Development (BILD) cohort (https://www.bild-cohort.ch/), unselected term and preterm infants were recruited to participate in this study [6]. Prematurity was defined as gestational age at birth < 37 weeks [7]. 16 (11 term and 5 preterm) infants (mean age 47 days, range 29-95 days, 50% girls), provided a total of 25 measurements, of which 9 were replicate measurements. Anthropometric data for infants are given in ***Table S3***.

*Breath collection*

Offline breath collection from infants was performed in unsedated sleep at 44 weeks of postconceptional age. The sampling device consisted of the following elements: medical compressed air (~21 % oxygen, ~78 % nitrogen), manometer, bypass flow, infant mask (face mask size 1, GaeleMed Corporation), t-piece (Hudson RCI®), one-way valve and a Nalophan® bag of approximately 500 mL with 450 $cm^2$ surface (***Figure 1***). Medical compressed air was administered with a flow rate of 0.3 L/min, helping infants to overcome the resistance of the device [8]. Excess air leaked through the bypass flow to the room. Infant face mask had a dead volume of 25 mL and T-pieces 30 mL. Infants' oxygen saturation and pulse were continuously monitored during breath collection with a pulse oximeter (Masimo Signal

Extraction Technology, Masimo). Offline breath collection was performed at room temperature until bags were completely filled (approximately 1 min). Flow rate of the infant breathing was not controlled for. If infants were still asleep after the first bag was filled (9 cases), a second bag was used to gather information on short-term repeatability. Bags were then transported to the SESI-HRMS laboratory within 10 min. The bags were manually deflated into the SESI source. The mass flow controller of the SuperSESI ion source was set at 0.7 L/min (identical value as for real-time analysis), which ensured a constant flow during the sample introduction procedure.

*Lung function and FeNO measurements*

Tidal breathing flow volume loops were captured at 44 weeks postconceptional age using Exhalyzer D (EcoMedics, Duernten, Switzerland) and FeNO measurement recorded by a chemiluminescence analyzer (CLD 77 AM; EcoMedics AG, Duernten, Switzerland) following ERS/ATS guidelines [9]. For analysis, we used the first 20-30 regular breaths during non-REM (non-rapid eye movement) sleep from the total recorded breathing. We excluded sighs and 10 breaths before and after a sigh. Simultaneous to tidal breathing recording, the fraction of exhaled nitric oxide (FeNO) was measured online with a chemiluminescence analyzer during the third quartile of expiration and averaged over the 20-30 breaths used for analysis[10]. Following ERS/ATS guidelines for infant lung function testing, mean tidal flows, volume and flow-volume loop were calculated. We investigated respiratory rate (RR), mean tidal volume ($V_T$) and minute ventilation ($V_E$). Ratio of time to peak tidal expiratory flow (PTEF) and expiratory time ($t_{PTEF}/t_E$) were used to describe TBFVL shapes.

*Data analysis*

The data matrix collected for the infants consisted of 25 samples × 2,284 features, whereby nine of the samples were short-term replicates. This initial matrix was visualized using t-distributed stochastic neighbor embedding (t-SNE) [11]. After computing the mean of the two short term replicates, the data matrix therefore consisted of 16 samples (unique individuals) × 2,284 features. Pearson's correlation coefficient between minute ventilation, FeNO and exhaled metabolites was computed.

**Sample size calculation**

The required sample size for a future study (***Figure 3d***) was determined based on Monte Carlo simulations generating 10,000 samples of size 5 to 500 to test the correlation between the two variables: 4-hydroxynonenal (4-HNE) and other clinical variables (e.g., minute ventilation). For a given sample size, we generated Monte Carlo simulations to determine an approximate cutoff value for a test of the correlation. We then generated samples under the alternative hypothesis, and estimated the power of the test, being:

Null hypothesis: exhaled 4-HNE and lung function variables are uncorrelated.

Alternative hypothesis: exhaled 4-HNE and lung function variables are correlated with an absolute value of *r* at least as high as 0.22 (4-HNE vs FeNO).

**Table S1** Anthropometric data from adult participants

| ID | Age (year) | Gender | Number of measurements |
|---|---|---|---|
| **DOPAEx2_BS07** | 36 | Female | 56 |
| **DOPAEx2_BS09** | 24 | Male | 50 |
| **DOPAEx2_BS10** | 28 | Female | 54 |
| **DOPAEx2_BS11** | 36 | Male | 20 |
| **DOPAEx2_BS12** | 36 | Male | 22 |
| **DOPAEx2_BS13** | 28 | Female | 22 |
| **DOPAEx2_BS14** | 32 | Female | 20 |
| **DOPAEx2_BS15** | 34 | Female | 22 |
| **DOPAEx2_BS16** | 30 | Female | 20 |
| **DOPAEx2_BS17** | 36 | Male | 22 |
| **DOPAEx2_BS18** | 43 | Female | 20 |
| **DOPAEx2_BS19** | 38 | Female | 22 |
| **DOPAEx2_BS20** | 38 | Male | 22 |

**Table S2** Anthropometric data from infant participants

| ID | Number of measurements | Gender | Birth | Age (days) | Respiratory rate (breaths/minute) | Mean tidal volume (mL) | Mean tidal volume/kg (mL/kg) | Minute ventilation (mL/minute) | Minute ventilation/kg (mL/minute-kg) | FeNO [ppb] |
|---|---|---|---|---|---|---|---|---|---|---|
| **BILD_4373** | 1 | Female | Term | 32 | | | | | | |
| **BILD_4374** | 1 | Male | Term | 33 | | | | | | |
| **BILD_4375** | 2 | Female | Term | 34 | 52 | 37 | 6.9 | 1933 | 358 | 12.4 |
| **BILD_4376** | 2 | Male | Term | 29 | 37 | 33 | 6.9 | 1248 | 260 | 25.8 |
| **BILD_4377** | 1 | Male | Term | 47 | 33 | 39 | 9 | 1299 | 302 | 17.9 |
| **BILD_4378** | 2 | Male | Term | 35 | 40 | 42 | 9.1 | 1670 | 363 | 13.1 |
| **BILD_4379** | 2 | Female | Term | 34 | | | | | | |
| **BILD_4381** | 2 | Female | Term | 35 | 41 | 32 | 8 | 1306 | 329 | 23.7 |
| **BILD_4382** | 2 | Female | Term | 35 | | | | | | |
| **BILD_4383** | 2 | Male | Preterm | 78 | 52 | 33 | 6 | 1729 | 332 | 18.2 |
| **BILD_4384** | 2 | Male | Term | 33 | | | | | | |
| **BILD_4385** | 2 | Female | Preterm | 41 | 29 | 36 | 9 | 1046 | 274 | 21.2 |
| **BILD_4386** | 1 | Male | Preterm | 75 | 39 | 38 | 8.5 | 1479 | 329 | 22.9 |
| **BILD_4387** | 1 | Male | Preterm | 75 | 37 | 33 | 7.4 | 1217 | 272 | 34.1 |
| **BILD_4389** | 1 | Female | Term | 33 | 35 | 38 | 8 | 1291 | 284 | 20.1 |
| **BILD_4390** | 1 | Female | Preterm | 95 | 39 | 41 | 8.5 | 1593 | 332 | 16.9 |
| | | | **Mean (STD)** | **47 (21)** | **39 (7)** | **37 (3)** | **8 (1)** | **1437 (266)** | **312 (36)** | **21 (6)** |

**Table S3** Compounds contained in standard gas mixture used for quality control

| MS polarity | m/z $[M+H]^+$ | Name | Formula |
|---|---|---|---|
| Positive | 59.0491414 | Acetone | $C_3H_6O$ |
| Positive | 69.0698769 | Isoprene | $C_5H_8$ |
| Positive | 73.0647915 | 2-Butanone | $C_4H_8O$ |
| Positive | 87.0804415 | 2-Pentanone | $C_5H_{10}O$ |
| Positive | 93.0698769 | Toluene | $C_7H_8$ |
| Positive | 105.0698769 | Styrene | $C_8H_8$ |
| Positive | 121.101177 | Mesitylene | $C_9H_{12}$ |
| Positive | 137.1324771 | α-Terpinene | $C_{10}H_{16}$ |

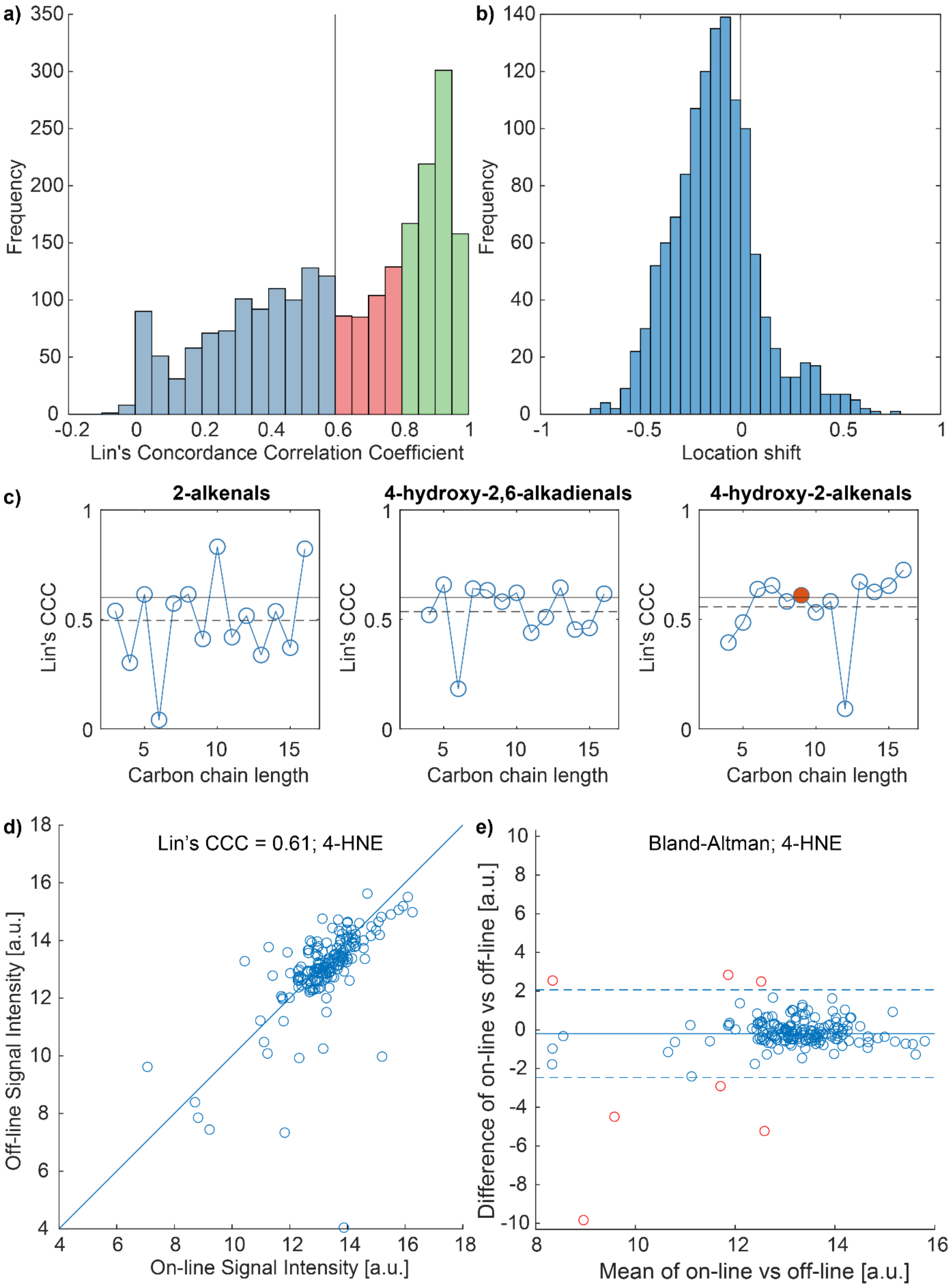


**Figure S1 a**) Distribution of Lin's CCC for the 2,284 measured features. **b**) Distribution of the location shift for the features with a Lin's CCC > 0.6 **c**) Lin's CCC vs. chain length for three homologous series of aldehydes: 2-alkenals, 4-hydroxy-2,6-alkadienals and 4-hydroxy-2-alkenals (red = 4-HNE; solid lines Lin's CCC = 0.6; dashed lines mean Lin's CCC for the group) **d**) Real-time vs offline signal intensity for all pairs of measurements for 4-HNE **e**) Bland-Altman plot shows the mean difference of 4-HNE between real-time and offline measurement pairs.

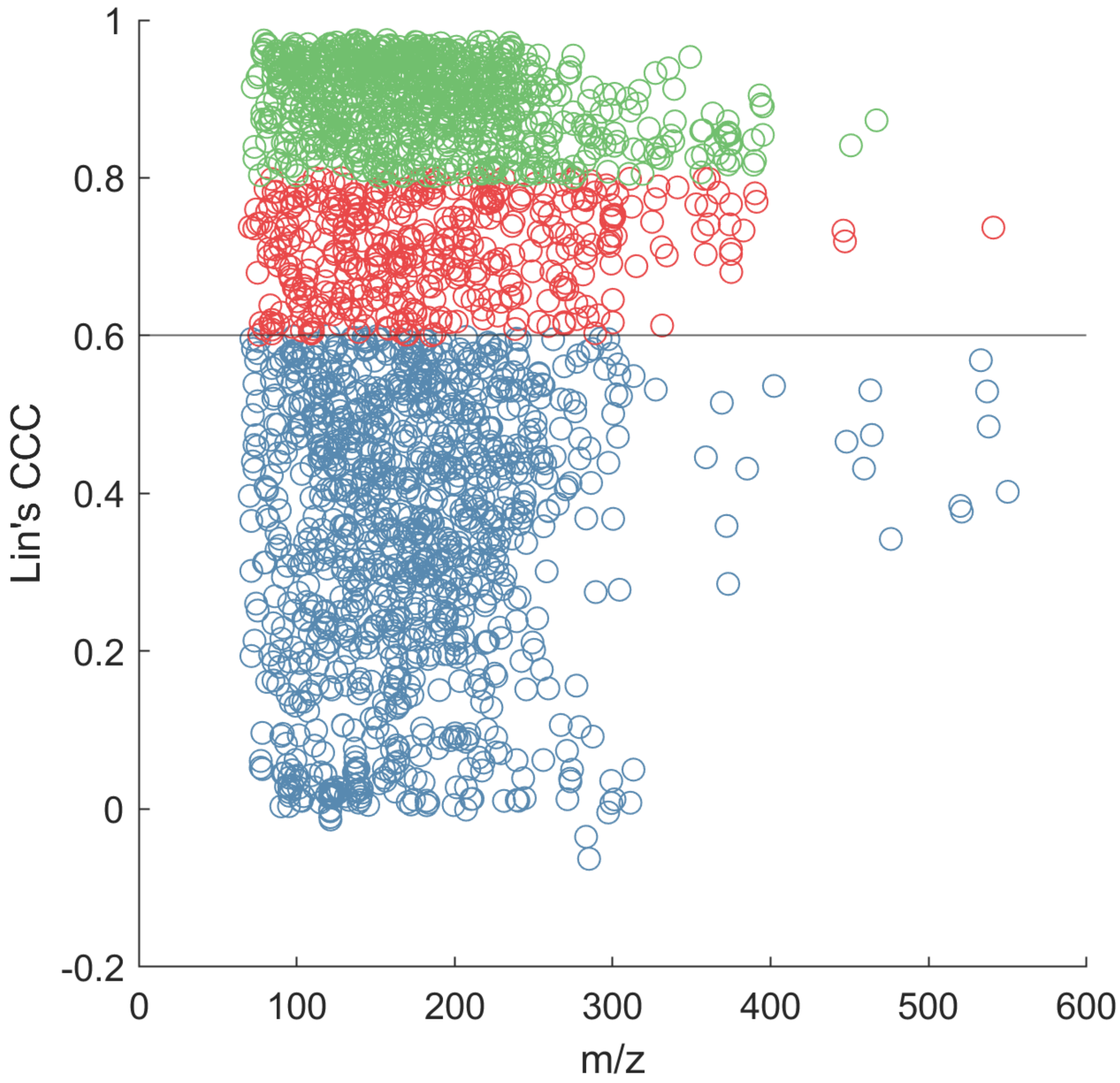


**Figure S2** Lin's concordance correlation coefficient as a function of m/z for all measurements of adults. Highest quality data points are colored in green (Lin’s CCC > 0.8), moderate in red (Lin’s CCC 0.6-0.8) and poor quality in blue (Lin’s CCC < 0.6)

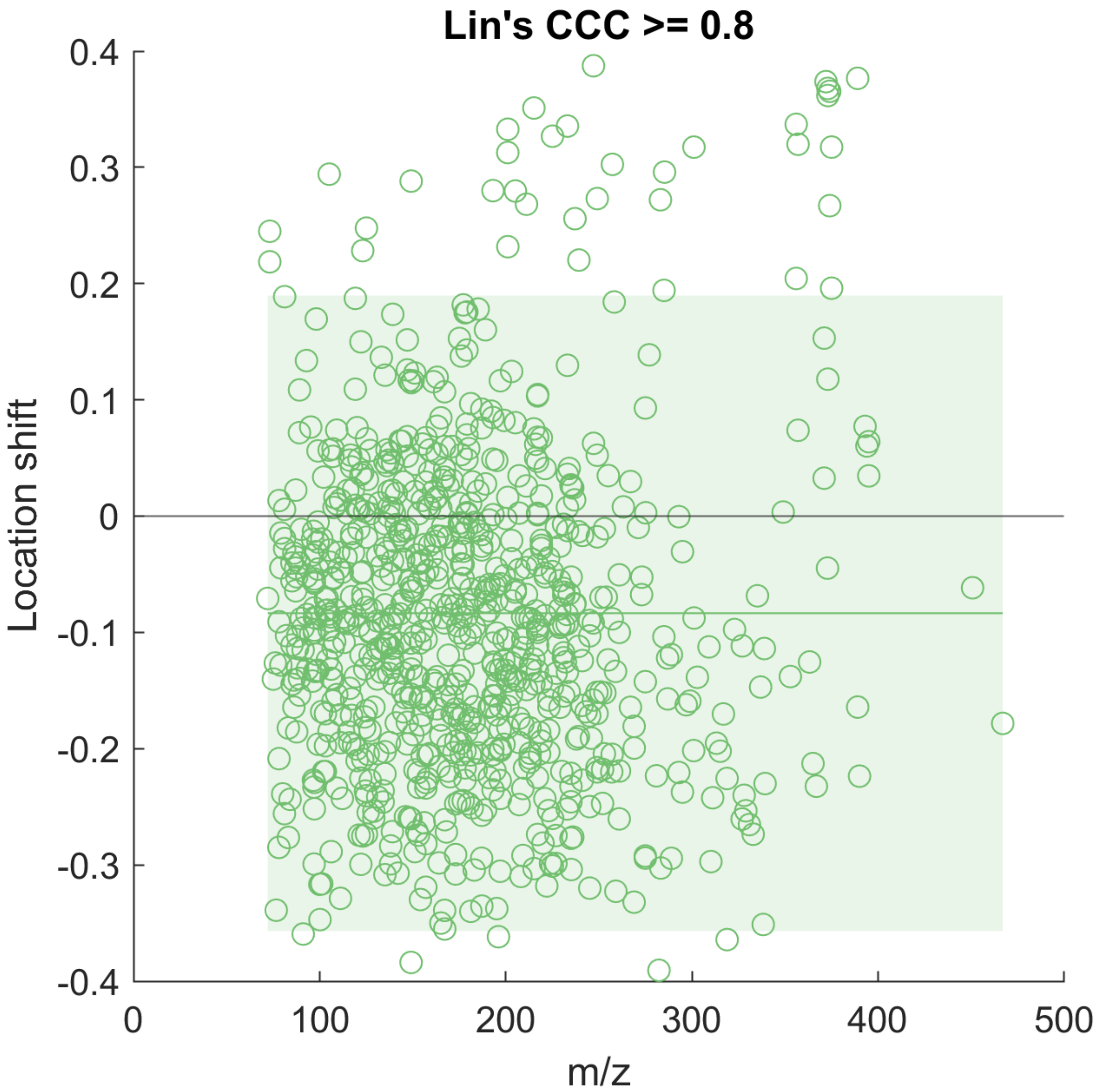


**Figure S3** Location shift as a function of m/z for measurements of adults with Lin's CCC ≥ 0.8. Green line indicates the mean and band represents 95% confidence interval.

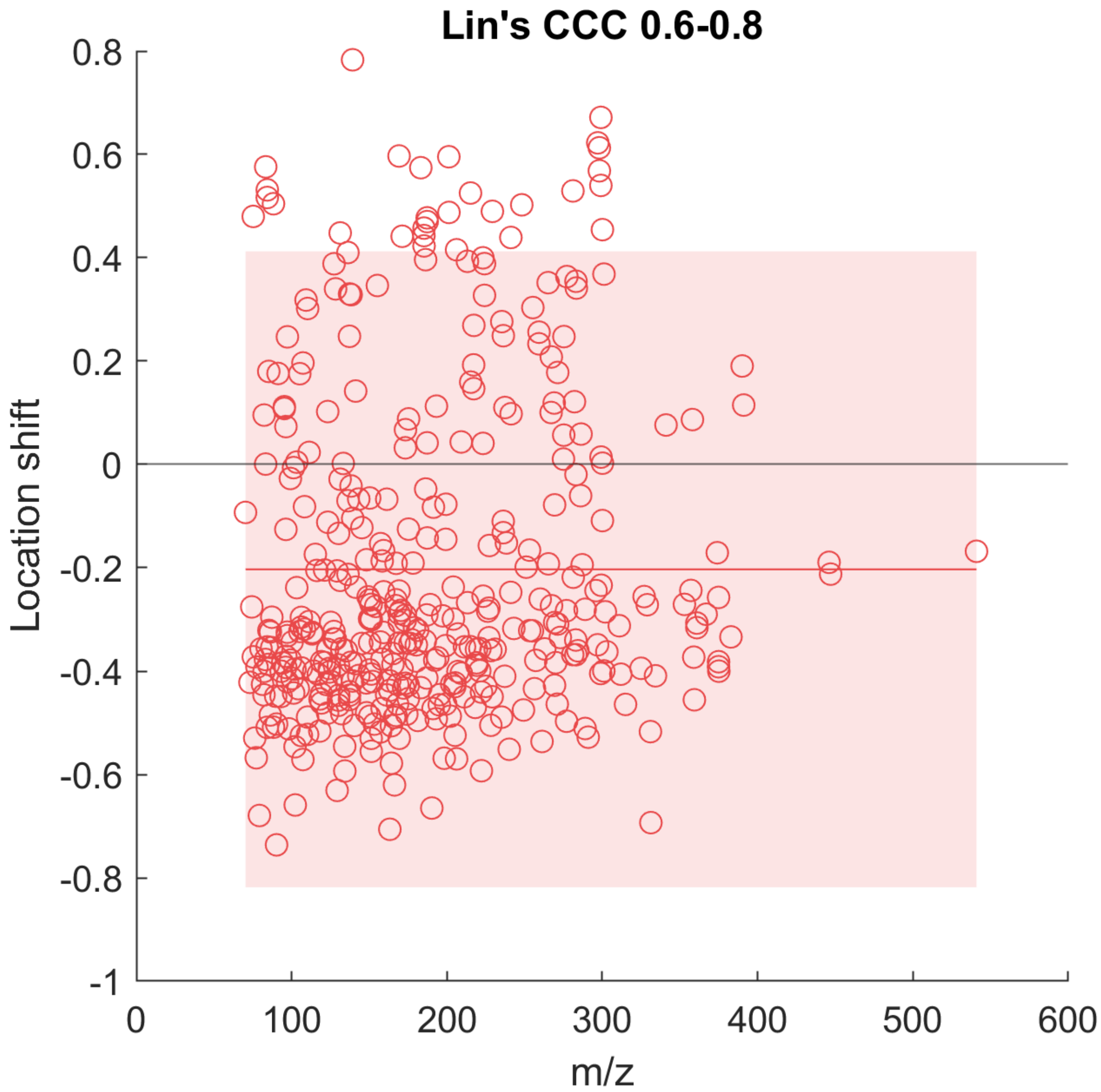


**Figure S4** Location shift as a function of m/z for measurements of adults with Lin's CCC ≥ 0.6 and < 0.8. Red line indicates the mean and band represents 95% confidence interval.

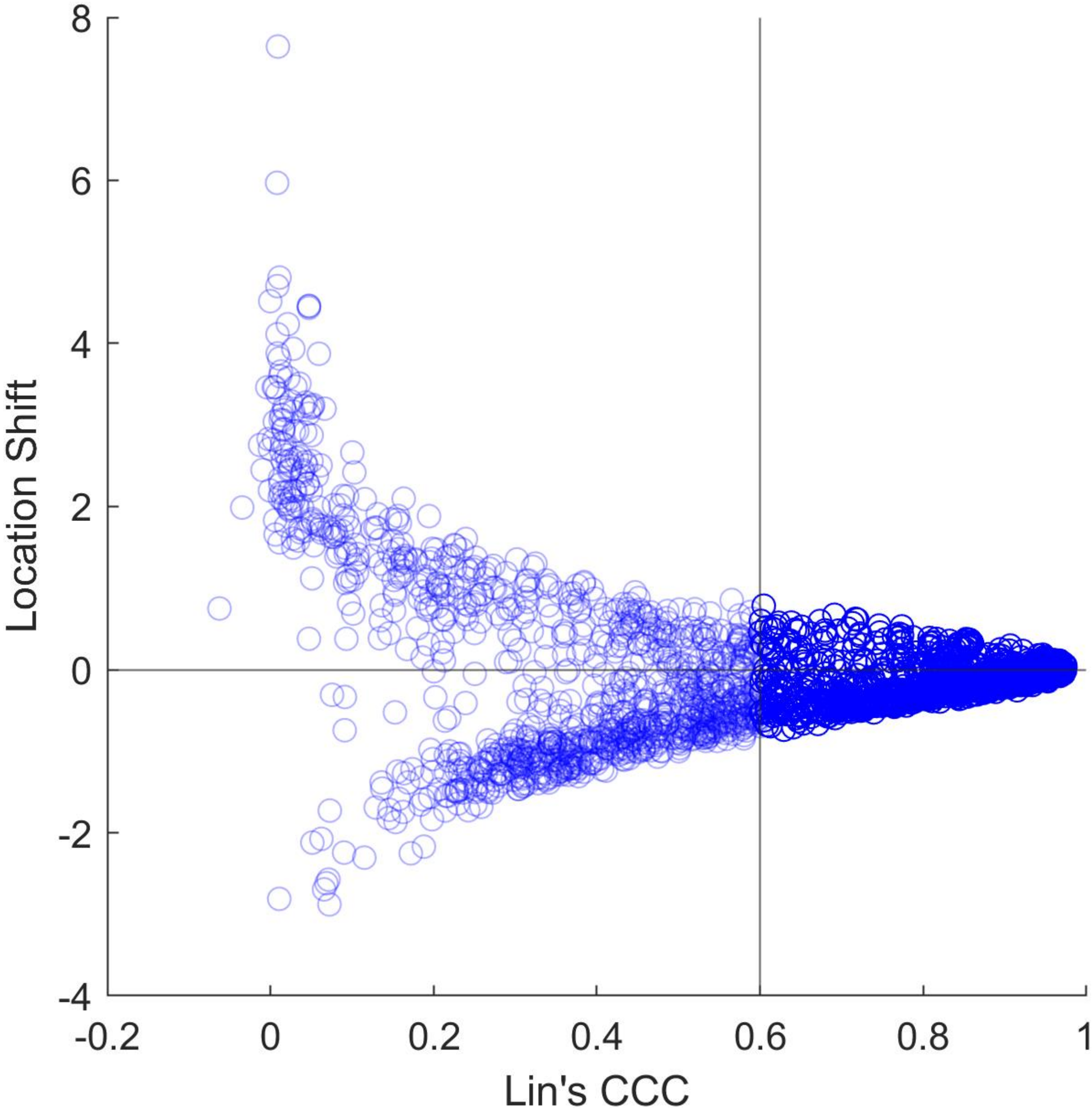


**Figure S5** Location shift as a function of Lin's concordance correlation coefficient**.**

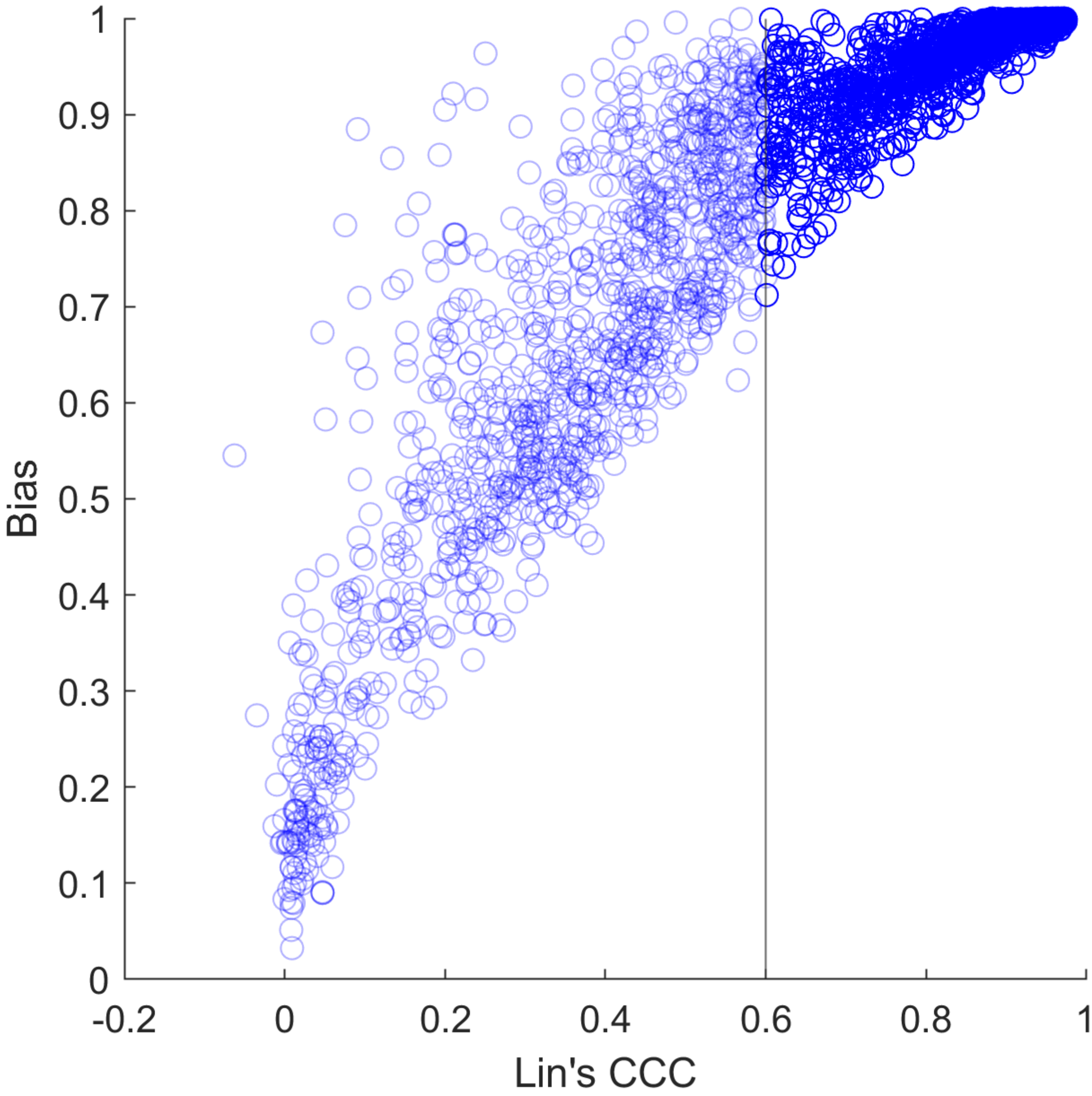


**Figure S6** Bias as a function Lin’s concordance correlation coefficient**.**

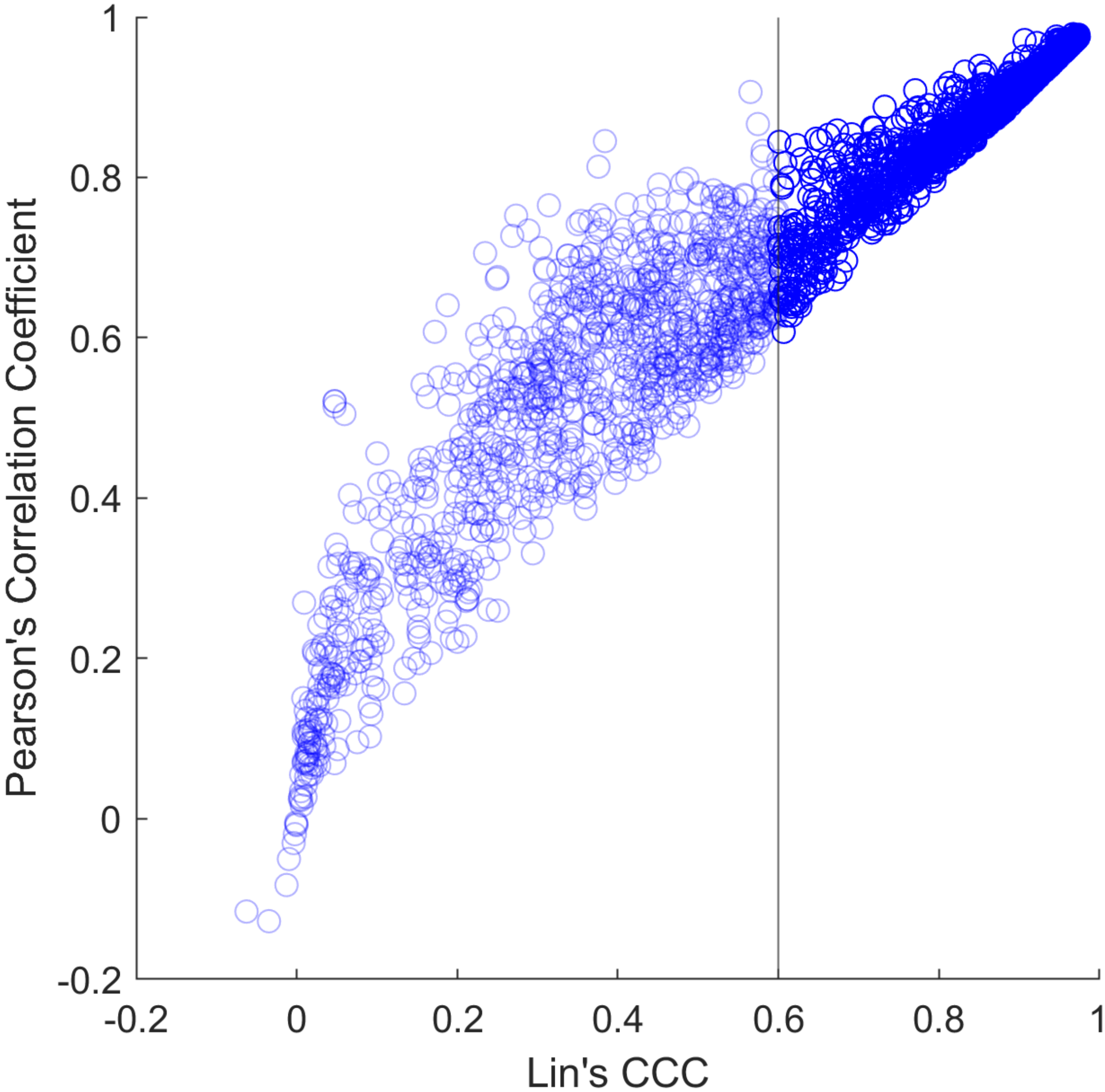


**Figure S7** Pearson’s correlation as a function of Lin’s concordance correlation coefficient

**References**


1 Gisler, A. *et al.* Real-time breath analysis of exhaled compounds upon peppermint oil ingestion by secondary electrospray ionization-high resolution mass spectrometry: technical aspects. *J. Breath Res.* **14**, 046001, doi:ARTN 046001

10.1088/1752-7163/ab9f8b (2020).

2 Boshier, P. R., Priest, O. H., Hanna, G. B. & Marczin, N. Influence of respiratory variables on the on-line detection of exhaled trace gases by PTR-MS. *Thorax* **66**, 919-920, doi:10.1136/thx.2011.161208 (2011).

3 Singh, K. D. T., Georgi; Decrue, Fabienne; Usemann, Jakob; Appenzeller, Rhea; Barreiro, Pedro; Jauma, Gabriel; Macia Santiago, Miriam; Vidal de Miguel, Guillermo; Frey, Urs; Sinues, Pablo. Standardization Procedures for Real-Time Breath Analysis by Secondary Electrospray Ionization-High Resolution Mass Spectrometry. *Anal Bioanal Chem* (2019).

4 Gisler, A. *et al.* Real-time breath analysis of exhaled compounds upon peppermint oil ingestion by secondary electrospray ionization-high resolution mass spectrometry: technical aspects. *Journal of breath research* **14**, 046001, doi:10.1088/1752-7163/ab9f8b (2020).

5 Lin, L. I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255-268 (1989).

6 Fuchs, O., Latzin, P., Kuehni, C. E. & Frey, U. Cohort profile: the Bern infant lung development cohort. *Int J Epidemiol* **41**, 366-376, doi:10.1093/ije/dyq239 (2012).

7 World Health Organization (WHO). *International Statistical Classification of Diseases and Related Health Problems; ICD-10*, <https://icd.who.int/browse10/2019/en#/P07.3> (2019).

8 Boshier, P. R., Cushnir, J. R., Priest, O. H., Marczin, N. & Hanna, G. B. Variation in the levels of volatile trace gases within three hospital environments: implications for clinical breath testing. *Journal of breath research* **4**, 031001, doi:10.1088/1752-7155/4/3/031001 (2010).

9 Bates, J. H., Schmalisch, G., Filbrun, D. & Stocks, J. Tidal breath analysis for infant pulmonary function testing. ERS/ATS Task Force on Standards for Infant Respiratory Function Testing. European Respiratory Society/American Thoracic Society. *The European respiratory journal* **16**, 1180-1192 (2000).

10 Fuchs, O. *et al.* Normative data for lung function and exhaled nitric oxide in unsedated healthy infants. *The European respiratory journal* **37**, 1208-1216, doi:10.1183/09031936.00125510 (2011).

11 van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J Mach Learn Res* **9**, 2579-2605 (2008).