Electronic Supplementary Material (ESI) for Analyst.

Supplementary Information

Title: **Comparison of breath sampling methods: a post hoc analysis from observational cohort studies**

Authors: Amalia Z. Berna, Chad L. Schaber, Lucy B. Bollinger, Mwawi Mwale, Rachel Mlotha-Mitole, Indi Trehan, Audrey R. Odom John

Correspondence to: aodom@wustl.edu

Methods

Patient enrolment:

Inclusion criteria: Children 3-15 years of age who were able to cooperate with breath collection. Children who had both a positive *P. falciparum* malaria RDT and blood smear were classified as having malaria, while those with both a negative rapid diagnostic test and blood smear were enrolled as controls. Children were included in the study if they reported fever (>38°C) at the time of enrolment.

Exclusion Criteria: Severe or cerebral malaria requiring urgent intervention, such as anti-epileptic medication, parenteral artesunate, respiratory support, or urgent blood transfusion. Receipt of any antimalarial therapy within the prior week of enrolment. Scented lotion or perfume use on the day of the study. Known diabetes mellitus, acute or chronic kidney or liver diseases.

After informed consent was obtained from caretakers, vital signs and anthropometry were taken, and a brief demographic and health history form was completed. Data was entered into a database and height-for-age and BMI-for-age Z scores were calculated using Anthro or Anthro Plus software as appropriate for age (World Health Organization, Switzerland).

Following consent from a parent or legal guardian, breath samples were collected as described in the Methods section of the mansucript. After specimen collection study participants returned to usual care as per the recommendations of the treating clinician. Antimalarial medications were provided for participants with positive malaria rapid diagnostic test results.

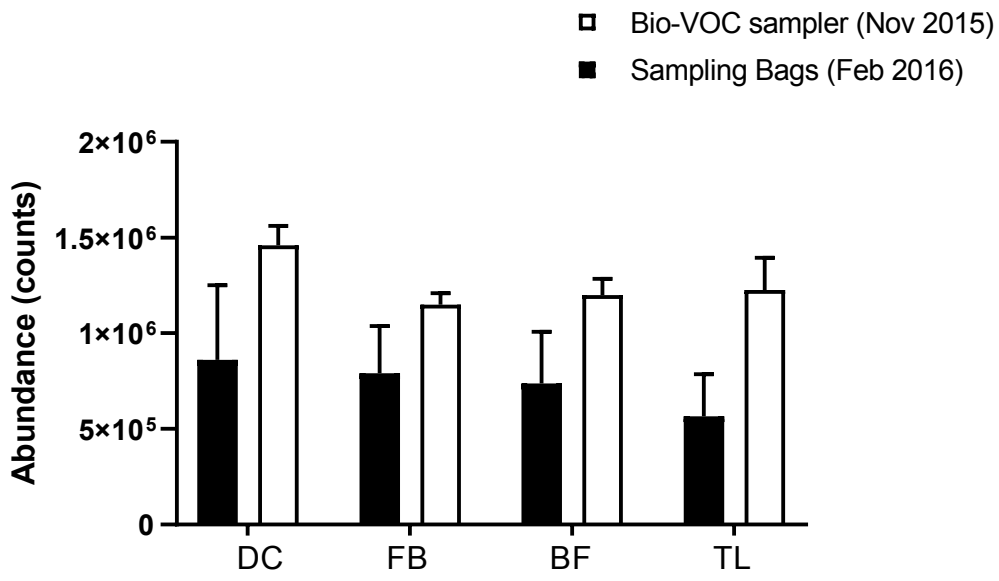
Additional information about the cohorts

The two cohorts did not differ significantly with respect to potential confounding clinical criteria (Supplementary Table 2), with the sole exception of the percentage of children reporting abdominal pain (64% in the Bio-VOC™ study; 86% in sampling bag study; p = 0.04), a finding of unclear significance.

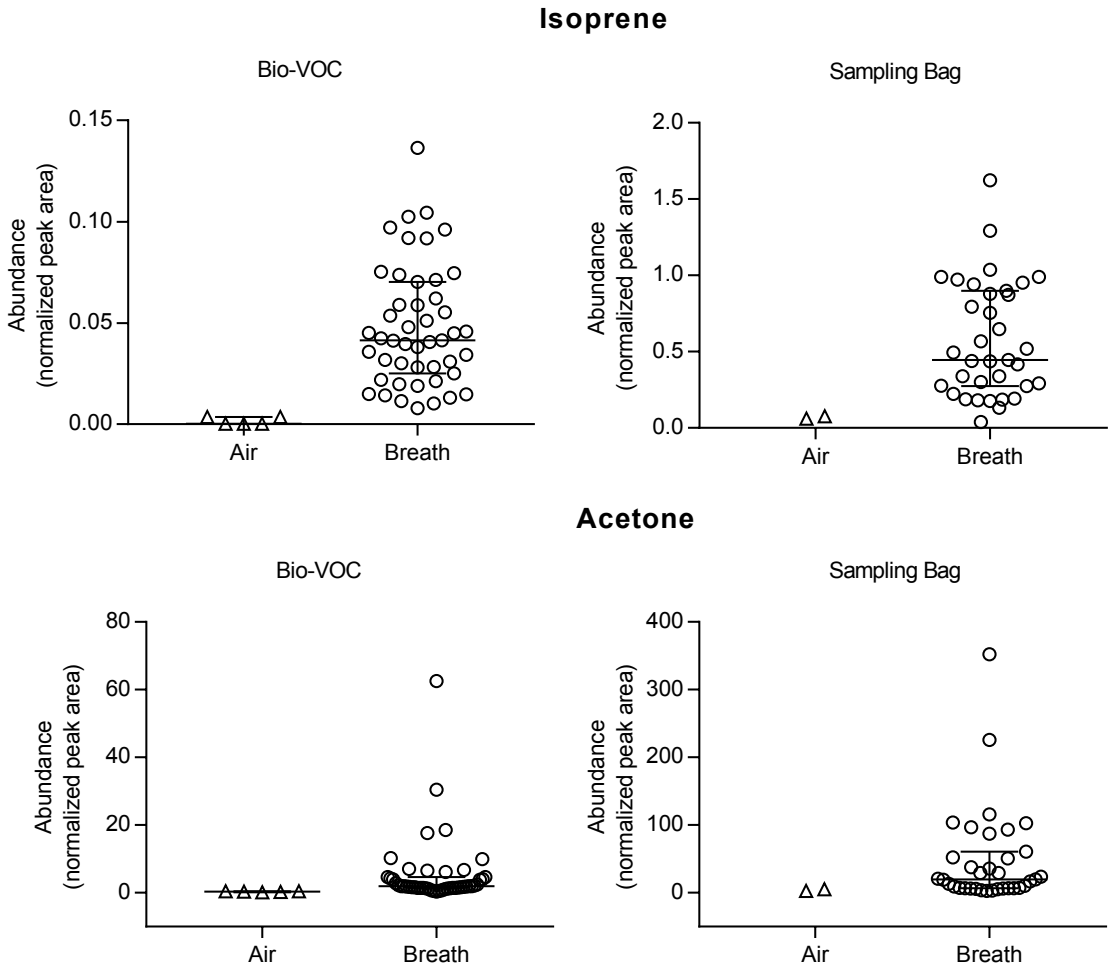
Breath VOCs were captured onto identical Universal sorbent tubes: inert stainless steel packed with Tenax 60/80, Carbograph 1 60/80, and Carboxen 1003 40/60 (Camsco, Texas). Prior to sampling, sorbent tubes were conditioned by flushing with 120 mL/min He at 290°C for one hour, or with 100 mL/min He at 320°C for two hours. Since samples were transported internationally (from Malawi to USA) for analysis, all breath samples were stored at -20°C prior to analysis, in order to preserve volatile integrity[1].

Sampling devices tests

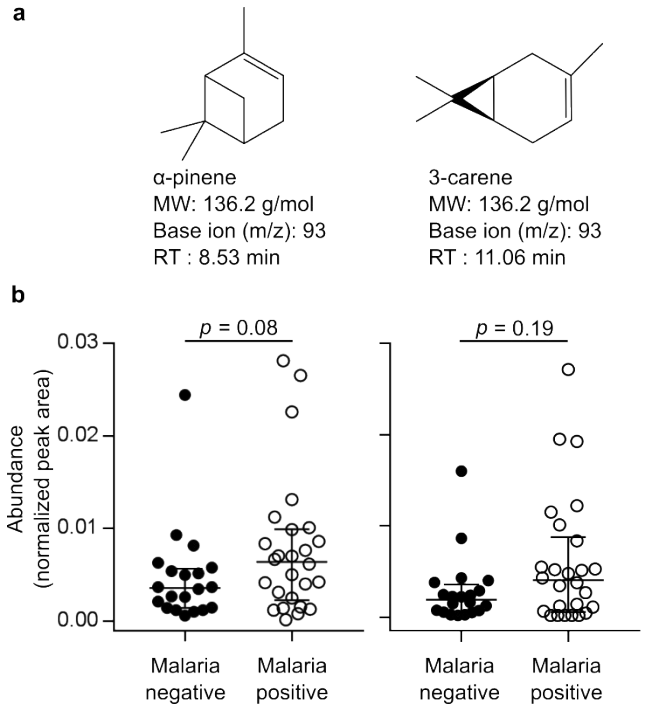
To identify contaminants arising from the polymer sampling bags, 3 new bags were filled with 1200 ml of high purity air (corresponding to 20.5% oxygen in nitrogen) and the air was transferred to sorbent tubes in the same manner as per breath samples (described in Experimental Section). For the Bio-VOC™ testing, the device was filled with high purity air and three volumes of the Bio-VOC™ were transferred to sorbent tubes, as per protocols for breath samples. Three repetitions were performed. The GC-MS analyses of these samples were carried out as previously detailed in the Experimental section for breath samples. Contaminants emitted by the Flexifilm bags and Bio-VOC™ are shown in Supplementary Table 3.



Supplementary Figure 1. Comparison of abundance of internal standards over the period of sample GC-MS analysis. The levels of internal standards (represented as mean and standard deviation) over two periods of GC/MS analysis: Nov 2015 (Bio-VOC™ sampler) and Feb 2016 (sampling bags). Injected concentration of standards were the same in both cohorts yet higher levels of the IS were observed in Nov 2015. The internal standards are: 1,2-Dichlorobenzene-d4 (DC), Fluorobenzene (FB), 4-Bromofluorobenzene (BF) and Toluene-d8 (TL).



Supplementary Figure 2. Isoprene (top figures) and acetone (bottom figures) are significantly more abundant in breath samples versus room air. Higher levels of acetone and isoprene were seen in breath versus room air for both the Bio-VOC™ (left figures) study and the sampling bag (right figures) study. Median and interquartile range are shown.



Supplementary Figure 3. Breath abundance of candidate mosquito-attractant terpenes in Bio-VOC™ study. Structure and chemical information of the terpenes. MW = molecular weight, RT = retention time. b) Breath abundance of terpenes α-pinene, left, and 3-carene, right, in children without (n = 20) and with (n = 26) falciparum malaria. Median and interquartile range are shown. For subjects in which compounds were not detected, abundance values were adjusted to the limit-of-quantification (0.0002). P-values, Mann Whitney U-tests.

Supplementary Table 1: Comparison of sample processing between Bio-VOC™ and Sampling bags

Description	Bio-VOC™ Study (*n* = 47)	Sample Bag Study (*n* = 35)
Age, median years (IQR)	8 (6-10)	8 (5-10)
Study location	Kamuzu Central Hospital and Bwaila Health Centre (Lilongwe, Malawi)	Kamuzu Central Hospital and Bwaila Health Centre (Lilongwe, Malawi)
Date of breath collection	November 2015	February 2016
Malaria season	High wet season	High wet season
Breath type collected	Alveolar air	Mixed expiratory air
Number of exhalations/sample	3	1-5
Volume of breath collected	264 - 387 mL	1 L
Enrolment period	2 weeks	2 weeks
Sample storage time, from collection to analysis	1 month	1 month
Pre-concentration method	Universal sorbent tubes	Universal sorbent tubes
GC-MS system	Pegasus 4D GCxGC-TOFMS system (LECO, Michigan)	Pegasus 4D GCxGC-TOFMS system (LECO, Michigan)

Supplementary Table 2. Comparison of patient demographic and clinical characteristics between the two study populations.

	Bio-VOC™ Study (*n* = 47)	**Sample Bag Study (*n* = 35)**	***p* value**[1]
Malaria Positive, *n* (%)	26 (55)	17 (49)	0.66
Demographics			
Age, median years (IQR)	8 (6-10)	8 (5-10)	0.86
Female, *n* (%)	20 (43)	18/34 (53)	0.38
Reported Symptoms, *n* (%)			
Fever	43 (91)	31 (89)	0.72
Diarrhoea	6 (13)	2 (6)	0.46
Vomiting	19 (40)	9 (26)	0.24
Headache	35 (74)	30 (86)	0.28
Abdominal Pain	30 (64)	30 (86)	0.04
Muscle/Joint Pain	23 (49)	16 (46)	0.83
Other, *n* (%)			
Chronic Malnutrition[2]	9/46 (20)	8/34 (24)	0.78
Acute Malnutrition[2]	3/46 (7)	1 (3)	0.63
Uses Bednet	29 (62)	19 (54)	0.65
Malaria within past 3 months	9/45 (20)	8/34 (24)	0.79

Data represented as number (%) except for age. If one or more patients were excluded due to gaps in the record, number given is fraction of total. Abbreviation: IQR, interquartile range. Data for Sampling Bag Study reported previously [16].

[1] Fisher's exact test or Mann-Whitney U-test used as appropriate to calculate *p* values.

[2] Chronic and acute malnutrition defined respectively as height-for-age Z-score or BMI-for-age Z-score two or more standard deviations below median.

Supplementary Table 3. Contaminants emitted by Flexfilm bags and Bio-VOC™. Compounds are ordered with respect to increasing retention time. "—" denotes that the VOC was not detected

Volatile organic compound	Quantifying ion	Average peak area Bio-VOC™	Average peak area Sample Bag
Methanesulfonyl chloride	79	9.84E+05	9.62E+05
Methyl methacrylate	69	—	1.78E+06
Toluene	91	8.52E+05	8.60E+05
2-Propanol, 1-(diethylamino)-	86	8.07E+04	—
Heptane, 2,4-dimethyl-	85	—	2.83E+06
3-Furfural	95	—	1.98E+05
1H-Pyrazole, 1-ethyl-3,5-dimethyl-	95	2.40E+05	—
Ethylbenzene	91	8.46E+04	1.43E+05
Octane, 4-methyl	43	—	3.42E+06
Phenylethyne	102	1.60E+05	6.55E+04
Styrene	104	2.34E+05	5.77E+05
Cyclopentanone, 2-methyl-	98	—	5.64E+05
Heptanal	70	2.68E+05	2.45E+05
Benzaldehyde	105	2.17E+06	2.49E+06
Benzene, 1,2,3-trimethyl-	105	—	1.06E+05
Benzofuran	118	1.33E+05	—
Bicyclo[2.2.1]heptane, 2-methyl-	69	2.54E+05	5.83E+05
Benzene, 1,4-dichloro-	146	1.97E+04	—

Pyridine, 4-ethyl-	79	5.71E+04	3.73E+04
Pentane, 2-cyclopropyl-	69	7.01E+05	—
Phenylethyl alcohol	91	3.10E+05	2.57E+05
Acetophenone	105	8.05E+05	1.36E+06
α-Cumyl alcohol	121	—	2.95E+05
Naphthalene	128	2.79E+05	4.68E+05
Unknown	67	4.65E+05	9.22E+05
Dodecane, 2,6,11-trimethyl-	71	—	2.47E+06

Text in bold are contaminants found in both devices

Reference:

1. M. Gjolstad, K. Bergemalm-Rynell, G. Ljungkvist, S. Thorud and P. Molander, *J. Sep. Sci.*, 2004, **27**, 1531-1539.