Bond, A., Greenwood, R., Lewis, S., Corfe, B., Sarkar, S., O'Toole, P., ... Probert, C. (2019). Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer. *Alimentary Pharmacology and Therapeutics*, *49*(8), 1005-1012. https://doi.org/10.1111/apt.15140

Received: 2 October 2018 | First decision: 23 October 2018 | Accepted: 20 December 2018

DOI: 10.1111/apt.15140

WILEY AP&T Alimentary Pharmacology & Therapeutics

Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer

Ashley Bond[1] | Rosemary Greenwood[2] | Stephen Lewis[3] | Bernard Corfe[4,5] | Sanchoy Sarkar[1] | Paul O'Toole[1] | Paul Rooney[1] | Michael Burkitt[1,6] | Georgina Hold[7] | Chris Probert[1,6]

[1]Royal Liverpool and Broadgreen University Hospital Trust, Liverpool, UK

[2]Research Design Service, School of Social and Community Medicine, University of Bristol, Bristol, UK

[3]Derriford Hospital, Plymouth, UK

[4]Molecular Gastroenterology Research Group, Department of Oncology, University of Sheffield, Sheffield, UK

[5]Insigneo Institute for in silico Medicine, University of Sheffield, Sheffield, UK

[6]Gastroenterology Research Unit, Department of Molecular and Cellular Physiology, Institute of Translational Medicine, University of Liverpool, Liverpool, UK

[7]Microbiome Research Centre, St George & Sutherland Clinical School, UNSW, Sydney, NSW, Australia

Correspondence
Prof. Chris Probert, Institute of Translational Medicine, University of Liverpool, Liverpool, UK.
Email: chris.probert@liverpool.ac.uk

Funding information
Ashley Bond received funding from the Shire Innovation Fund to fund the purchase of laboratory consumables

Summary

Background: Colorectal cancer remains a leading cause of mortality and morbidity. The UK Bowel Cancer Screening Programme (BCSP) has demonstrated that detection of colorectal cancer at an earlier stage and identification of advanced pre-malignant adenomas reduces mortality and morbidity.

Aim: To assess the utility of volatile organic compounds as a biomarker for colorectal neoplasia.

Methods: Faeces were collected from symptomatic patients and people participating in the UK BCSP, prior to colonoscopy. Headspace extraction followed by gas chromatography mass spectrometry was performed on faeces to identify volatile organic compounds. Logistic regression modelling and 10-fold cross-validation were used to test potential biomarkers.

Results: One hundred and thirty-seven participants were included (mean age 64 years [range 22-85], 54% were male): 60 had no neoplasia, 56 had adenomatous polyp(s) and 21 had adenocarcinoma. Propan-2-ol was significantly more abundant in the cancer samples ($P < 0.0001$, $q = 0.004$) with an area under ROC (AUROC) curve of 0.76. When combined with 3-methylbutanoic acid the AUROC curve was 0.82, sensitivity 87.9% (95% CI 0.87-0.99) and specificity 84.6% (95% CI 0.65-1.0). Logistic regression analysis using the presence/absence of specific volatile organic compounds, identified a three volatile organic compound panel (propan-2-ol, hexan-2-one and ethyl 3-methyl- butanoate) to have an AUROC of 0.73, with a person six times more likely to have cancer if all three volatile organic compounds were present ($P < 0.0001$).

Conclusions: Volatile organic compound analysis may have a superior diagnostic ability for the identification of colorectal adenocarcinoma, when compared to other faecal biomarkers, including those currently employed in UK.

Clinical trial details: National Research Ethics Service Committee South West - Central Bristol (REC reference 14/SW/1162) with R&D approval from University of Liverpool and Broadgreen University Hospital Trust (UoL 001098).

The Handling Editor for this article was Professor Peter Gibson, and it was accepted for publication after full peer-review.

1 | INTRODUCTION

Colorectal cancer is a leading cause of mortality and morbidity worldwide, with an estimated European incidence of 43.5 per 100 000 in 2012 and mortality of 19.5 per 100 000.[1] The lifetime risk, for UK residents, is 1 in 15 for men or 1 in 19 for women.[2] Across Europe, colorectal cancer is the second most common cause of cancer-related mortality.[1] Colorectal cancer carries a significant financial burden for the National Health Service, with a mean annual cost of £12 000 and £8800 for each patient diagnosed with rectal and nonrectal colon cancer respectively.[3] Data from the UK Bowel Cancer Screening Programme have clearly demonstrated that detection of colorectal cancer at an earlier stage and identification of advanced pre-malignant adenomas can reduce future cancer-associated mortality and morbidity.[4,5]

The UK Bowel Cancer Screening Programme uses a faeces-based screening tool to select patients to take forward to colonoscopy, in line with European guidance.[6] Currently, in England, the guaiac-based faecal occult blood testing (gFOBt) is employed. This test relies on bleeding from neoplastic lesions and can be used to identify people with >10 mL rectal blood loss daily. gFOBt is however, prone to false positive results after ingestion of certain foods.[7] The low sensitivity of gFOBt has led to criticism of its use for population-based screening.[8] The gFOBt is likely to be replaced by faecal immunochemical testing (FIT). FIT detects twice as many advanced cancers as guaiac testing [9] and can provide both qualitative and quantitative results. A recent observational study, from Italy, demonstrated a reduction in colorectal cancer-related mortality in regions where screening with FIT was adopted compared with regions where screening had not yet been implemented.[10,11] Burch et al[4] reported a meta-analysis of 59 studies of FOBT: sensitivities for the detection of all neoplasms ranged from 6.2% to 83.3% for gFOBTs and 5.4% to 62.6% for FITs, depending on the preferred specificity.[12] A review by NICE concluded that FIT has a specificity ranging from 43% to 86%.[13] However, FIT has limitations: the Dutch colorectal cancer screening programme reported 77% sensitivity with FIT based on 18 716 samples (specificity was not reported) and 23% of the patients developed interval cancers.[14]

Several studies have reported volatile organic compounds emitted from different substrates as biomarkers for colorectal cancer. One such study used selected ion flow tube mass spectrometry (SIFT-MS) to detect volatile organic compounds in faeces.[15] Another analysed urine, from patients with colorectal cancer, employing Field Asymmetric Ion Mobility Spectrometer (FAIMS).[16] The third used breath analysed by thermal-desorber gas chromatography–mass spectrometry (GCMS) in an attempt to diagnose colorectal cancer.[17,18] These were mainly proof of concept or feasibility studies that reported output patterns rather than identifying the individual compounds. Therefore, understanding the biological plausibility for patterns of volatile organic compounds can be difficult to interpret.

We undertook a prospective study of the volatile organic compounds emitted from faecal samples obtained from patients at risk of colorectal cancer.

2 | MATERIALS AND METHODS

2.1 | Participants

Most participants were recruited from colonoscopy waiting lists at the Royal Liverpool University Hospital (n = 122). Participants were referred by the Merseyside and Wirral Bowel Cancer Screening Programme with positive FOBt or patients undergoing colonoscopy for adenomatous polyp surveillance, planned polypectomy, the investigation of iron deficiency anaemia (IDA), change in bowel habit or abnormal radiological imaging. All patients recruited via the Bowel Cancer Screening Programme had a prior positive gFOBt. The FOBt status of the non- Bowel Cancer Screening Programme patients was unknown. No patients were assessed by FIT. Patient referrals and Bowel Cancer Screening Programme referrals were vetted to assess suitability and all consecutive patients were sent collection kits in the post. A subset of the faecal samples was provided from a cohort of symptomatic patients undergoing colonoscopy in Sheffield and Plymouth, UK.

Research ethics committee approval for the study was obtained from the National Research Ethics Service Committee South West - Central Bristol (REC reference 14/SW/1162) with R&D approval from University of Liverpool and Broadgreen University Hospital Trust (UoL 001098) from where patients were recruited over a 12-month period. All patients were supplied with an information sheet and provided written consent. Specific permission was also granted by the NHS Bowel Cancer Screening Programme Research Committee. Samples collected from Sheffield (n = 11) and Plymouth (n = 6) were acquired in line with existing ethical approval (North Sheffield Research Ethics Committee (Ref: 06/Q2308/93 and 13/SW/0238, respectively.

2.2 | Sample collection and storage

Samples were produced, at home, during the 48 hours preceding their colonoscopy and before commencing the required bowel preparation. The stool was produced initially into a foil dish, then participants were asked to place at least three spoonfuls of faeces into a glass vial (OdoReader, University of the West of England), before it was sealed and stored in a cool place, either outside or in the fridge. The initial volume of stool supplied by the patient was not specified but could not exceed the volume of the provided 20 mL glass vial. The sample was brought to the Endoscopy Department when the patients attended for the colonoscopy. During the transportation from the patient's home to the hospital the sample would have been at ambient temperature. Patients who had received antibiotics in the preceding 3-6 months and vegetarians were excluded. Colonoscopy results, including any histological findings, were recorded. Patients were categorised as having no neoplasia, adenomatous polyp(s) or cancer. Therefore, control patients were those with no neoplasia, but they could have had other abnormalities including diverticulosis. Patients with active colitis were excluded. The location, size and number of polyps were recorded. Polyps were assigned to the

adenoma group only after histological confirmation. Hyperplastic polyps were classified as no neoplasia. Demographic details, smoking status and antibiotic use were also recorded.

2.3 | Headspace volatile organic compound analysis

Four hundred and fifty milligram of unadulterated faeces was aliquoted into new 10 mL headspace vials and sealed with magnetic caps (Supleco, Poole, UK).[19] Both the sample intended for analysis and the residual faeces were then stored at −20°C until GCMS analysis was performed.

Headspace volatile organic compounds analysis was performed using a Combipal (CTC, Zwingen, Switzerland) and carboxen/polydimethylsiloxane solid phase microextraction fibre (Sigma Aldrich, Dorset, UK). The fibre was exposed to the headspace above the faeces for 20 minutes. Volatile organic compounds were analysed by GCMS (Perkin Elmer Clarus 500 quadrupole, Beaconsfield, UK): volatile organic compounds were thermally desorbed from the fibre at 220°C in the injection port of the GCMS for 5 minutes. Injection was made in splitless mode and a split of 50 mL/min was turned on 2 minutes into the run. Helium carrier gas of 99.996% purity (BOC, Guildford, UK) was passed through a helium purification system, Excelasorb™ (Supelco) at 1 mL/min. The GC column was a 60 metre long Zebron ZB-624 capillary column with an inner diameter of 0.25 mm. The column (Phenomenex, Macclesfield, UK) was lined with a 1.4 μm film of 94% dimethyl polysiloxane and 6% cyanopropylphenyl. The GCMS temperature program of the run was as follows: initial oven temperature was held at 40°C for 2 minutes then the temperature was ramped up at a rate of 5°C/min to 220°C, with a 4 minute hold at this temperature to give a total run time of 42 minutes. The mass spectrometer was run in electron impact (EI) ionisation mode, scanning the mass ion range 10-300 at 0.05 scan/s. A 4 minute solvent delay was used at the start of the run.[19–21]

2.4 | Data processing

The GCMS data were processed using a pipeline involving the Automated Mass Spectral Deconvolution and Identification System software (AMDIS, Version 2.71, 2012), the NIST mass spectral library (version 2.0, 2011) and the R (R core team, 2013) package Metab.[22] AMDIS and NIST software were used to build a volatile organic compound library containing 162 metabolites present in the stool samples analysed in this study. A forward and reverse match of 800/1000 and above was used for assigning tentative compound identifications. Using this volatile organic compound library, AMDIS was then applied to deconvolute chromatograms and identifying metabolites. The report generated by AMDIS was further processed by Metab, in order to align metabolites and recalculate their relative abundances based on the intensity of a specific ion mass fragment per metabolite. In order to develop robust parsimonious statistical models, those compounds found to be present in fewer than 20% of the patients in both groups were removed.[20,21] Compounds were named using IUPAC nomenclature.

2.5 | Statistical analysis

Data analysis was performed in R, Stata and Metaboanalyst,[23] utilising Student's *t* test, Mann-Whitey tests, Fisher's exact test, ANOVA, false discovery rate correction, Partial Least Squared Discriminant Analysis (PLS-DA), factor analysis and Receiver Operator Characteristic (ROC) analysis. Logistic regression modelling, along with 10-fold cross-validation was used to test potential biomarkers. When Metaboanalyst was used the data were normalised by median and log-transformed. When Mann-Whitney and factor analysis was used the data were logged, normalised and the absence of a volatile organic compound substituted by the value -3 to create an artificial floor in keeping with the concept that the lack of an observable volatile organic compound is analogous to the least amount measurable.

3 | RESULTS

One hundred and thirty-seven patients were included in the study: the average age was 64.3 years; 56% were male. The mean age was lowest in those with no neoplasia and greatest in those with the cancer, $P = 0.02$. None of the participants reported being smokers or vegetarians. Self-reported ethnicity was noted: all but one was White British. 27.7% of study participants were recruited from the Bowel Cancer Screening Programme.

One hundred and sixty-two volatile organic compounds were identified in whole sample set. The mean number of volatile organic compounds identified in each group was similar: cancer (mean 54.3, standard deviation [SD] 1.2), adenoma (mean 55.0, SD 11.6) or controls (mean 54, SD 10.3). Biomarker identification focused on higher risk neoplastic disease, namely established colorectal cancer and >4 individual polyps of any size.

Initially samples from patients in all three groups were compared using ANOVA. Fourteen volatile organic compounds differed in abundance: after adjusting for multiple comparisons, none were significant, but several were of interest as they were found in later comparisons, including 5-methyl-2-propan-2-yl-cyclohexan-1-ol, ethyl 3-methylbutanoate and propan-2-ol (Table 1).

3.1 | Volatile organic compounds as a biomarker for colonic adenocarcinoma: quantitative analysis

PLS-DA comparing those with no neoplasia and those with colorectal cancer showed a separation that suggested potential diagnostic utility (Figure 1). Exploration of potential candidates for biomarker analysis can be seen in Table 2. These comparisons did not include samples from patients with adenomatous polyps: only those with confirmed adenocarcinoma and no neoplasia were included for analysis.

Propan-2-ol and 5-methyl-2-propan-2-yl-cyclohexan-1-ol was further considered in isolation, following assessment when combining volatile organic compound as a ratio. The latter was formerly known

TABLE 1 Demographic and clinical features of participants recruited in Liverpool, Sheffield and Plymouth

	No-neoplasia	Adenoma	Cancer
Number	60	56	21
Mean age, years (range)	61.9 (22-85)	65.6 (41-84)	72.7 (64-78)
Gender			
Male	25	36	7
Female	34	20	3
Smoker (yes)	0	0	0
Indication for colonoscopy			
BCSP	13	22	3
IDA	16	6	5
Change in bowel habit-diarrhoea	11	4	1
Surveillance previous neoplasia/FH	10	24	1
IBD assessment/ surveillance	9	0	0
GI bleeding	1	0	0
Unknown	0	0	11

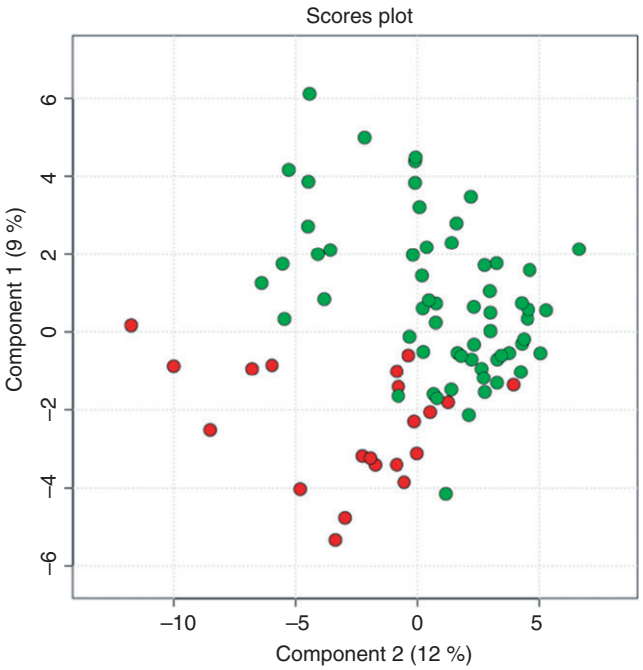

FIGURE 1 Partial least square discriminate analysis comparing those with adenocarcinoma of the colon (red) and no colonic neoplasia (green)

as dl-menthol: we will use that name to aid readability. Propan-2-ol selected as it was the volatile organic compound most strongly associated with cancer; dl-menthol as it was the only volatile organic compound to be negatively associated with cancer.

The abundance of propan-2-ol was compared in the three groups using Kruskal Wallis test. The mean abundance in cancer was 88.7×10^6, in adenoma 23.7×10^6 and controls 51.5×10^6; the differences were significant, $P = 0.001$ (Figure 2). The data were log-transformed and compared using ANOVA: the differences were significant ($P = 0.01$), post hoc Dunnett testing showed the main difference was between samples from patients with cancer and controls ($P = 0.007$): this implies that, while the mean for adenomas was appeared less than that for controls, the adenoma data were widely spread. It is noteworthy, of the other compounds associated with cancer, three are esters of propan-2-ol with short chain acids.

TABLE 2 Volatile organic compounds with abundance that differs significantly between samples from patients with cancer and controls

VOC	*P* value	*q* value	Association with CRC:
Propan-2-ol[a]	<0.0001	0.004	Positive
Hexan-2-one	0.01	0.77	Positive
Ethyl 3-methylbutanoate[a]	0.03	0.77	Positive
Propan-2-yl butanoate	0.03	0.77	Positive
Propan-2-yl pentanoate	0.03	0.77	Positive
1,4-xylene	0.03	0.77	Positive
Propan-2-yl propanoate	0.04	0.77	Positive
5-methyl-2-propan-2-yl-cyclohexan-1-ol[a]	0.05	0.77	Negative

Significance was determined by Student's *t* test applied to log-transformed data. The *q* value was reported after adjustment for multiple comparisons.

[a]Volatile organic compounds identified in the ANOVA.

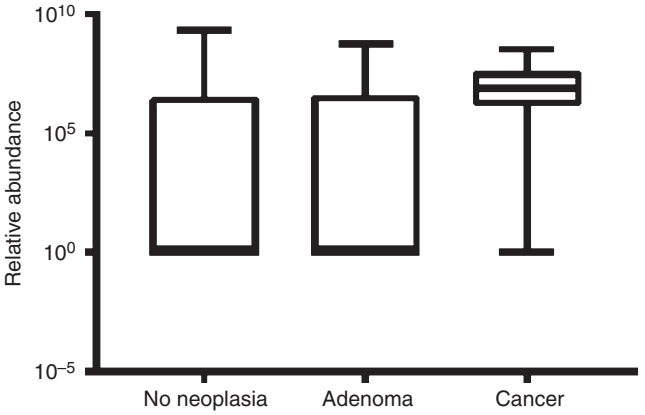

FIGURE 2 Box plots to show the relative abundance of propan-2-ol in faeces from all participants. All patients in each cohort are included Normal (no neoplasia) n = 60, adenoma n = 56 and cancer n = 21

The abundance of dl-menthol was subjected to the same analysis. The mean abundance in cancer was 0.7×10^6, in adenoma 15.1×10^6 and controls 8.3×10^6; the differences were significant, $P = 0.04$. The data were log-transformed and compared using ANOVA: the differences were significant ($P = 0.003$), post hoc Dunnett testing showed patients with cancer had significantly less dl-menthol than adenoma and control groups.

Propan-2-ol showed the most promise as a single biomarker for colorectal cancer: it achieved an area under the ROC (AUROC) curve

(Figure 3) to predict colorectal cancer of 0.76 with a sensitivity of 83% and specificity of 71%.

Calculating ratios of all possible metabolite pairs and then choosing top ranked ratios, based on p values, allowed for further biomarker assessment.

A hold-out technique was applied to the 81 samples (21 cancer and 60 controls) in order to validate the combination of 3-methyl butanoic acid/propan-2-ol as a biomarker for colorectal cancer: 50% of each cohort were held back. The combination of 3-methylbutanoic acid and propan-2-ol gave the best result: data from patients with cancer and with no neoplasia were modelled using logistic regression and 10-fold cross-validation, based upon the abundance of 3-methylbutanoic acid and propan-2-ol (Table 3): AUROC is 0.86, sensitivity 87.9% (95% CI 0.87-0.99) and specificity 84.6% (95% CI 0.65-1.0).

3.2 | Assessing for patterns of volatile organic compounds as biomarkers for colonic adenocarcinoma- factor analysis using qualitative data

Principal component analysis and a non-orthogonal rotation feature analysis was applied to qualitative (presence/absence) data for volatile organic compounds using all volatile organic compounds that was present in at least 30% of the group for any of the three diagnostic groups. Using all the data, the solution could not be extracted due to convergence issues (because many of the volatile organic compounds were highly correlated with each other) until the number of extracted factors had been reduced from 19 to 17.

By looking at the factors, rather than the individual volatile organic compounds, to fit a regression model to predict cancer a number of different orthogonal rotations were used to produce a set of potential predictors. This process highlighted the combination of propan-2-ol, hexan-2-one and ethyl 3-methylbutanoate as a key predictor. Used as continuous variables directly extracted from the data set (prior to logging and normalisation) the simple summation of the quantities of these three peaks produces AUROCs of 0.768 and 0.750. Using a simple summation of the presence and absence of all three volatile organic compounds as a biomarker panel predicted cancer patients distinctly from all other patients with a $P = 0.001$ and an AUROC of 0.73 and predicted cancer versus normal with a $P = 0.006$ and an AUROC 0.702, suggesting very little information is lost by using just presence and absence of these three compounds. It is noteworthy that these three volatile organic compounds were also found by the univariate analysis, before correction for multiple comparisons.

Pure reference solutions of propan-2-ol, hexan-2-one and ethyl 3-methylbutanoate confirmed the identification within the stool samples was correct.

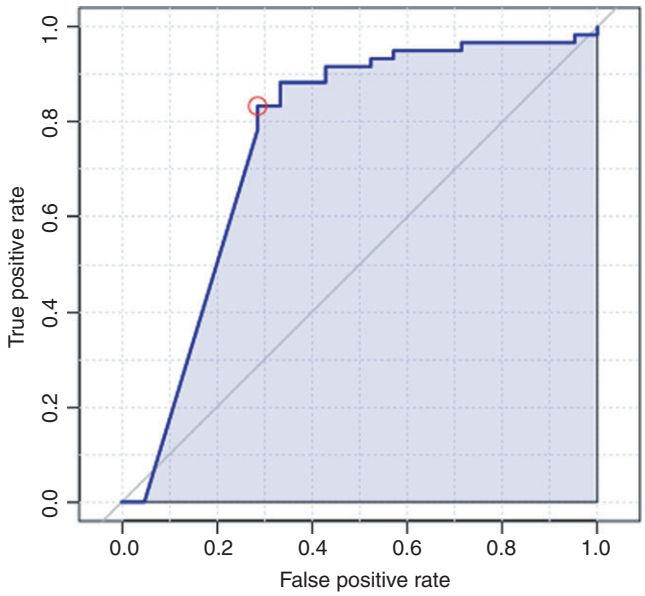

FIGURE 3 Receiver operating characteristic curve for propan-2-ol when comparing those with adenocarcinoma of the colon and no colonic neoplasia

TABLE 3 Area under the receiver operating characteristic results for the volatile organic compounds emitted when using a comparison of ratios for those with adenocarcinoma of the colon and no colonic neoplasia

VOC ratio combination	AUROC	95% CI	Optimal sensitivity %	Optimal specificity %
3-methylbutanoic acid/ propan-2-ol	0.82	0.71-0.92	81	76
dl-Menthol/propan-2-ol	0.82	0.7-0.91	85	71
Nonan-2-one/ propan-2-ol	0.81	0.7-0.91	88	57
2-methylpropanoic acid/propan-2-ol	0.81	0.7-0.91	80	76
Decane/propan-2-ol	0.81	0.68-0.9	81	71
Pentyl acetate/ propan-2-ol	0.8	0.69-0.89	83	71
Phenylacetaldehyde/ propan-2-ol	0.79	0.65-0.91	86	76

3.3 | Volatile organic compounds as a biomarker for colonic adenomas: quantitative analysis

Several volatile organic compounds were associated with samples from patients with >4 polyps (Figure 4). None of the associations remained after adjustment for multiple comparisons.

4 | DISCUSSION

Correctly identifying patients to undergo colonoscopy as part of population-based screening is vital in order to maximise pathology capture and to minimise unnecessary examinations. There is a clear link to improved outcomes from colorectal cancer through the identification of earlier stage colorectal cancer and pre-malignant

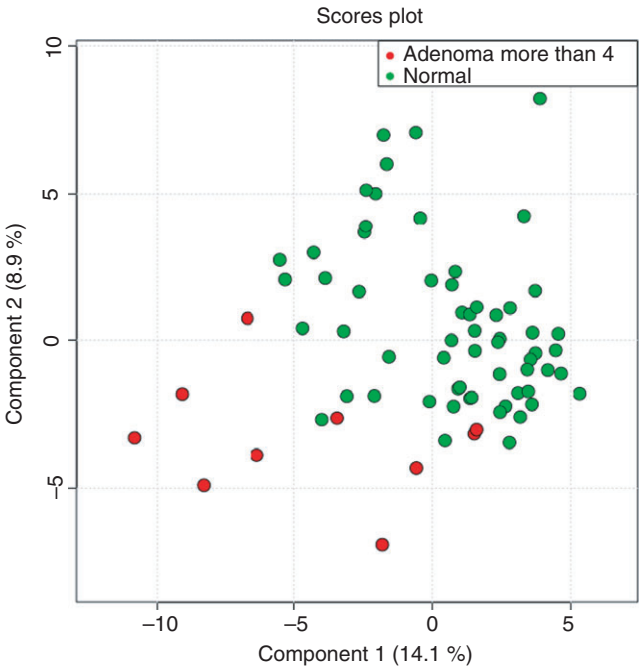

FIGURE 4 Partial least square discriminate analysis using all the identified VOCs, comparing those with no neoplasia against those with > 4 individual polyps of any size

adenomatous colonic polyps.[5] This study has demonstrated the utility of volatile organic compounds emitted from faeces to act as a biomarker for colonic neoplasia, in particular adenocarcinoma.

We have reported two volatile organic compound-based models for the identification of samples from patients with adenomas and colorectal cancer. In the quantitative approach, the models were dominated by the presence of propan-2-ol either as an alcohol or as an ester with short chain fatty acids. The qualitative model, which simply used presence or absence of compounds, also included propan-2-ol.

Propan-2-ol is a secondary alcohol that may be derived from acetone: a pathway associated with Clostridria.[24] The role of propan-2-ol in the pathogenesis of colorectal cancer had not been proposed before: the occurrence in this study may be a bystander phenomenon linked to dysbiosis; further work is needed. Ethyl 3-methylbutanote probably arises from a condensation reaction between ethanol and 3-methylbutanoic acid. Ethanol is produced by several metabolic pathways. 3-Methylbutanoic acid is derived from 3-methylbutanal, by aldehyde dehydrogenase: the aldehyde is derived from leucine.

Using a variety of methods and substrates, other studies have suggested a utility of volatile organic compound analysis for the diagnosis of GI disease,[20,25–27] including colorectal cancer. One such study, from 2015, used selected ion flow tube mass spectrometry to analysis volatile organic compounds emitted from faeces of FOBt positive patients. Comparing patients with no neoplasia and high grade neoplasia, ions probably arising from hydrogen sulphide, dimethyl sulphide and dimethyl disulphide were significantly higher in samples from high risk compared to low risk subjects. The authors reported overall specificity of 78% and 72% sensitivity (Table 4).[15] Two separate studies, from 2014 and 2013, reported the analysis of volatile organic compounds found in urine and breath, respectively. The study examining urine used Field Asymmetric Ion Mobility Spectrometer (FAIMS): 133 patients were included; 83 colorectal cancer patients and 50 healthy controls. Sensitivity and specificity for colorectal cancer detection with FAIMS were 88% and 60% respectively.[16] A third technology, in the form of thermal-desorber gas chromatography–mass spectrometry, was used to assess volatile organic compounds in the study examining breath. Assessing the pattern of 15 compounds showed a sensitivity of 86%, a specificity of 83% and AUROC of 0.85.[17] More recently, using the same technique, this group described the ability of exhaled volatile organic compounds to discriminate between colorectal cancer patients before and after curative surgery.[18] A further study from 2014 reported the utility of a pattern recognition–based detection technique, using volatile organic compounds found in faeces. This study did not attempt to identify the individual compounds but focused on differing patterns. It attempted to identify established colorectal cancer and pre-malignant adenomatous lesions. Faecal volatile organic compound profiles of patients with colorectal cancer differed significantly from controls (AUROC, 0.92; sensitivity, 0.85; and specificity, 0.87). Patients with advanced adenomas could also be distinguished from controls (AUROC, 0.79; sensitivity, 0.62; and specificity, 0.86).

Population-based screening or a point of care test is the most likely clinical application of such volatile organic compound analysis. Despite their relatively low patient acceptance rates, faecal based techniques are currently the most commonly employed ie, FOBt, either gFOBt or FIT. The gFOBT currently used in the UK Bowel Cancer Screening Programme has a sensitivity of 36% and a specificity of 94% for the detection of colorectal cancer.[28,29] To date, there are no controlled trials that demonstrate that FIT is superior to gFOBT or to no screening in terms of reducing colorectal cancer-related mortality in average risk persons. However, a recent observational study from Italy demonstrated a reduction in colorectal cancer-related mortality in regions where screening with FIT was adopted compared with regions where screening had not yet been implemented [10,11] The superiority of FIT over gFOBts is now widely recognised and the European Quality Assurance Guideline on

TABLE 4 Propan-2-ol, hexan-2-one and ethyl 3-methylbutanoate in stool from patients with colonic adenocarcinoma, adenomatous colonic polyps and no neoplasia

	Adenoma	Cancer	Normal
Mean number of these three VOCs	1.16	2.0	1.33
Proportion with none of these three VOCs	23.2% (13/56)	0% (0/21)	20.0% (12/60)
With just 1	42.8% (24/56)	23.8% (5/21)	36/7% (22/60)
With just 2	28.6% (16/56)	52.4% (11/21)	33.3% (20/60)
With all three	5.4% (3/56)	23.8% (5/21)	10.0% (6/60)

Colorectal Cancer Screening published in 2011 recommends FIT in preference to gFOBT.[30,31] Studies have reported FIT to have overall sensitivity for colorectal cancer was 0.79 (95% CI: 0.69-0.86) and the overall specificity was 0.94 (95% CI: 0.92-0.95).[32] Various countries have adopted FIT into their colorectal cancer screening programmes and the Bowel Cancer Screening Programme plans to replace gFOBt with FIT.[33] Comparing the result of our study it would appear that volatile organic compounds have a greater diagnostic ability than either FOBt for the identification of colorectal cancer. In the future patient acceptability may be improved by the use of ingestible capsules.[34,35]

Further work is necessary to ascertain the source of the volatile organic compounds that were found in association with colorectal cancer and adenomas. It is likely that they are bacterial metabolites. The driver-passenger model of colorectal cancer development suggests that *Fusobacterium nucleatum* is the key to ongoing tumourogenesis, with butanoic acid playing a key role in supporting the tumour microenvironment.[36] The presence of *F. nucleatum* in colorectal cancer tissue has also been noted in more advanced colorectal cancer, particularly those with lymph node metastasis, again supporting the positive correlation.[37,38] *F. nucleatum* (data not shown, paper in preparation) has been shown to produce propan-2-ol (data not shown, paper in preparation) and may be a source of propan-2-ol in colorectal cancer samples.

Moreover, we demonstrated a significant decrease in dl-menthol in those with colorectal cancer. This commonly originates from dental hygiene products. *F. nucleatum* is found in the oral cavity and thus poor dental hygiene is linked to increase in *F. nucleatum* and potentially increased the risk of colorectal cancer. Thus, the absence of dl-menthol might indicate the presence of poor hygiene and the carriage of *F. nucleatum*.

The heterogenous nature of the study cohort is a limitation as it limits the generalisability of the results to an asymptomatic screening population. As with techniques employed in population-based screening there is reliance on the patients to appropriately collect and handle the samples, our methods has this limitation, therefore potentially introducing error here. All attempts were made in the patient selection, sampling equipment, storage, transportation and laboratory analysis to minimise volatile organic compound contamination and variability. We wanted to simplify the procedures as much as possible in this pilot. Patients collected samples in their own homes and brought them to the Endoscopy Department just as they do for calprotectin assessment. Any influence of handling samples in this way would have acting upon cases and controls and is unlikely to have materially affected the statistical separation of the data.

5 | CONCLUSIONS

This pilot study has found compounds that are positively and negatively associated with the presence of colorectal neoplasia. Volatile organic compounds emitted from faeces can be utilised as a biomarker for colorectal cancer. Prospective studies are required to determine whether volatile organic compounds are better than FIT testing or whether they should be used together.

ACKNOWLEDGEMENT

Declaration of personal interests: None.

AUTHORSHIP

Guarantor of the article: Ashley Bond.

Author contributions: AB devised the study with CP, recruited patients, collected samples, conduced the laboratory work, data analysis and drafted the manuscript. RG provided statistical support and analysis. SL and BC provided additional samples and reviewed the manuscript prior to submission. SS, PR and POT facilitated patient recruitment and reviewed the manuscript prior to submission. MB assisted with laboratory work, provided statistical support and reviewed the manuscript prior to submission. GH performed laboratory work and reviewed the manuscript prior to submission. CP devised the study, oversaw its completion, assisted with drafting the manuscript and editing prior to submission.

All authors approved the final version of the manuscript.

ORCID

Ashley Bond https://orcid.org/0000-0002-3237-4782
Stephen Lewis https://orcid.org/0000-0003-3367-2949
Bernard Corfe https://orcid.org/0000-0003-0449-2228
Georgina Hold https://orcid.org/0000-0001-7573-3397
Chris Probert https://orcid.org/0000-0003-4550-0239

REFERENCES

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*. 2013;49:1374-1403.
2. Rees CJ, Bevan R. The National Health Service Bowel Cancer Screening Program: the early years. *Expert Rev Gastroenterol Hepatol*. 2013;7:421-437.
3. Trueman P, Lowson K, Chaplin S, et al. Bowel cancer services: costs and benefits. 2007.
4. Logan RF a, Patnick J, Nickerson C, et al. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. *Gut* 2012;61:1439-1446.
5. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med*. 2014;370:2541.
6. Altobelli E, Lattanzi A, Paduano R, et al. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. *Prev Med (Baltim)* 2014;62:132-141.
7. Colorectal cancer association of Canada. A guide to FOBT and FIT tests. Guaiac-based FOBT Immunochem. FOBT; 2016. http://www.colorectal-cancer.ca/en/screening/fobt-. Accessed July 1, 2018.
8. Tonus C, Neupert G, Sellinger M. Colorectal cancer screening by non-invasive metabolic biomarker fecal tumor M2-PK. *World J Gastroenterol*. 2006;12:7007-7011.

9. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. *Gut*. 2014;64:784-790.
10. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut*. 2015;64:1327-1337.
11. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. *Gut*. 2015;64:784-790.
12. Burch J, Soares-Weiser K, St John D, et al. Diagnostic accuracy of faecal occult blood tests used in screening for colorectal cancer: a systematic review. *J Med Screen*. 2007;14:132-137.
13. National Institute for Health and Care Excellence (NICE). Quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care. Diagnostics guidance [DG30]. 2017.
14. van der Vlugt M, Grobbee EJ, Bossuyt PMM, et al. Interval Colorectal Cancer Incidence Among Subjects Undergoing Multiple Rounds of Fecal Immunochemical Testing. *Gastroenterology*. 2017;153:439.e2-447.e2.
15. Batty CA, Cauchi M, Lourenço C, et al. Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer. *PLoS ONE*. 2015;10:e0130301.
16. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, et al. Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. *PLoS ONE*. 2014;9:e108750.
17. Altomare DF, Di Lena M, Porcelli F, et al. Exhaled volatile organic compounds identify patients with colorectal cancer. *Br J Surg*. 2013;100:144-150.
18. Altomare DF, Di Lena M, Porcelli F, et al. Effects of Curative Colorectal Cancer Surgery on Exhaled Volatile Organic Compounds and Potential Implications in Clinical Follow-up. *Ann Surg*. 2015;262:862-867.
19. Reade S, Mayor A, Aggio R, et al. Optimisation of sample preparation for direct SPME-GC-MS analysis of murine and human faecal volatile organic compounds for metabolomic studies. *J Anal Bioanal Tech* 2014;5:184.
20. Bond A, Vernon A, Reade S, et al. Investigation of volatile organic compounds emitted from faeces for the diagnosis of giardiasis. *J Gastrointest Liver Dis*. 2015;24:281-286.
21. Khalid T, Aggio R, White P, et al. Urinary Volatile Organic Compounds for the Detection of Prostate Cancer. *PLoS ONE*. 2015;10:e0143283.
22. Aggio R, Villas-Bôas SG, Ruggiero K. Metab: an R package for high-throughput analysis of metabolomics data generated by GC-MS. *Bioinformatics*. 2011;27:2316-2318.
23. Xia J, Sinelnikov I, Han B, Wishart D. MetaboAnalyst 3.0 - making metabolomics more meaningful. *Nucleic Acid Res* 2015;43:W251-W257.
24. Walther T, François JM. Microbial production of propanol. *Biotechnol Adv*. 2016;35:984-996.
25. Smolinska A, Bodelier AGL, Dallinga JW, et al. The potential of volatile organic compounds for the detection of active disease in patients with ulcerative colitis. *Aliment Pharmacol Ther*. 2017;45:1244-1254.
26. Baranska A, Mujagic Z, Smolinska A, et al. Volatile organic compounds in breath as markers for irritable bowel syndrome: a metabolomic approach. *Aliment Pharmacol Ther*. 2016;44:45-56.
27. Arasaradnam RP, Covington JA, Harmston C, et al. Review article: next generation diagnostic modalities in gastroenterology - Gas phase volatile compound biomarker detection. *Aliment Pharmacol Ther*. 2014;39:780-789.
28. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *N Engl J Med*. 2004;351:2704-2714.
29. Collins JF, Lieberman DA, Durbin TE, et al. Accuracy of screening for fecal occult blood on a single stool sample obtained by digital rectal examination: a comparison with recommended sampling practice. *Ann Intern Med* 2005;142:81-86.
30. Halloran SP, Launoy G, Zappa M. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition–Faecal occult blood testing. *Endoscopy* 2012;44:SE65-SE87.
31. Sung JJ, Ng SC, Chan FK, et al. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. *Gut*. 2015;64:121-132.
32. Song L-L, Li Y-M. Current noninvasive tests for colorectal cancer screening: an overview of colorectal cancer screening tests. *World J Gastrointest Oncol*. 2016;8:793.
33. Department of Health. Improving outcomes: a strategy for cancer. 2011.
34. Kalantar-Zadeh K, Ha N, Ou J, et al. Ingestible sensors. *ACS Sensors*. 2017;28:468-483.
35. Berean K, Ha N, Ou J, et al. The safety and sensitivity of a telemetric capsule to monitor gastrointestinal hydrogen production in vivo in healthy subjects: a pilot trial comparison to concurrent breath analysis. *Aliment Pharmacol Ther*. 2018;48:646-654.
36. Tjalsma H, Boleij A, Marchesi JR, et al. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat Publ Gr*. 2012;10:575-582.
37. Wu N, Yang X, Zhang R, et al. Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb Ecol*. 2013;66:462-470.
38. Castellarin M, Warren RL, Freeman JD, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res* 2012;22:299-306.

SUPPORTING INFORMATION

Additional supporting information will be found online in the Supporting Information section at the end of the article.

How to cite this article: Bond A, Greenwood R, Lewis S, et al. Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer. *Aliment Pharmacol Ther*. 2019;49:1005–1012. https://doi.org/10.1111/apt.15140