Irritable bowel syndrome and active inflammatory bowel disease diagnosed by faecal gas analysis.

Short title

Diagnosing IBD and IBS from faecal gas

Authors

Raphael B. M. Aggio¹, Paul White², Hiruni Jayasena³, Ben de Lacy Costello⁴, Norman M. Ratcliffe⁴ and Chris S. J. Probert¹

Affiliations

¹ Institute of Translational Medicine, Department of Cellular and Molecular Physiology, University of Liverpool, Liverpool, United Kingdom.

² Department of Engineering, Design and Mathematics, University of the West of England, Bristol, United Kingdom.

³ Department of Clinical Medicine, Faculty of Medicine, General Sir John Kotelawala Defence University, Ratmalana, Sir Lanka

⁴ Institute of Biosensor Technology, University of the West of England, Bristol, United Kingdom.

Funding

Part of this work was supported by a University Translational Award from the Wellcome Trust.

Word count: 4346

Corresponding author

Raphael Bastos Mareschi Aggio,

Department of Cellular and Molecular Physiology,

Institute of Translational Medicine,

University of Liverpool,

Nuffield Building,

Crown Street,

Liverpool,

United Kingdom

L693GE

Raphael.Aggio@liverpool.ac.uk

Phone: +44 1517954010

Fax: + 44 1517946825

Disclosure

This manuscript has not been previously published and the manuscript is not under consideration elsewhere. R.B.M. Aggio, P. White, H. Jayasena, B. de Lacy Costello, N.M. Ratcliffe, and C.S.J. Probert certify that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: B. de Lacy Costello, N.M. Ratcliffe and C.S.J. Probert are the inventors of the intellectual property related to applications of the gas chromatography-sensor. The intellectual property is owned by their employers, the University of Liverpool and the University of West of England. In addition, R.B.M. Aggio and C.S.J. Probert are the inventors of the intellectual property. The other authors have nothing to disclose.

Authorship Statement

Guarantor of article:

C.S.J. Probert

Author contribution:

R.B.M. Aggio performed the data analysis, wrote the manuscript and generated the figures; P. White performed the data analysis, wrote the manuscript and generated figures; H. Jayasena collected patient data and samples and revised the

manuscript; B. de Lacy Costello developed the gas chromatography-sensor system and revised the manuscript; N.M. Ratcliffe developed the gas chromatography-sensor system and wrote the manuscript; C.S.J. Probert developed the gas chromatography-sensor system and wrote the manuscript.

All authors approved the final version of this manuscript.

ABSTRACT

Background - Inflammatory bowel disease and irritable bowel syndrome may present in a similar manner. Measuring faecal calprotectin concentration is often recommended to rule out inflammatory bowel disease, however, there are no tests to positively diagnose irritable bowel syndrome and invasive tests are still used to rule out other pathologies. Aim – To investigate a platform technology for diagnosing inflammatory bowel disease and irritable bowel syndrome based on faecal gas. Methods - The platform technology is composed of a gas chromatography column coupled to a metal oxide gas sensor (OdoReader) and a computer algorithm. The OdoReader separates the volatile compounds from faecal gas and the computer algorithm identifies resistance patterns associated with specific medical conditions and builds classification models. This platform was applied to faecal samples from 152 patients: 33 patients with active inflammatory bowel disease; 50 patients with inactive inflammatory bowel disease; 28 patients with irritable bowel syndrome; and 41 healthy donors (Control). Results - The platform classified samples with accuracies from 75 to 100% using rigorous validation schemes namely leave-one-out cross validation, 10-fold cross validation, double-cross validation and their Monte Carlo variations. The most clinically important findings, after double-cross validation, were the accuracy of active Crohn's disease vs. irritable bowel syndrome (87%; C.I. 84-89%) and irritable bowel syndrome vs. controls (78%; C.I. 76-80%). These schemes provide an estimate of out-of-sample predictive accuracy for similar populations. **Conclusions -** This is the first description of an investigation for the positive diagnosis of irritable bowel syndrome and diagnosing inflammatory bowel disease.

Keywords

Inflammatory bowel disease; Irritable bowel syndrome; volatile compounds.

1 - Introduction

Irritable bowel syndrome is a chronic relapsing gastrointestinal disorder characterised by abdominal pain, bloating and a change in bowel habit¹. The disorder can be diagnosed on the symptoms alone, especially in younger patients and those with a long history. At present, the preferred means of diagnosing irritable bowel syndrome is the application of the Rome criteria². However, despite the recommendations of the American College of Gastroenterology³ and the British Society of Gastroenterology¹, many clinicians still view irritable bowel syndrome as a diagnosis of exclusion and perform numerous investigations to rule out organic diseases⁴.

Inflammatory bowel disease, ulcerative colitis and Crohn's disease, are also chronic relapsing gastrointestinal disorders and their symptoms may resemble irritable bowel syndrome. The use of faecal calprotectin in primary care is being promoted to aid referral to secondary care for patients suspected to have inflammatory bowel disease. In essence, the calprotectin test is being used to rule out organic diseases in the hope that the primary care physicians will manage patients with irritable bowel syndrome. Irritable bowel syndrome is more common than inflammatory bowel disease and it is not surprising that some patients develop irritable bowel syndrome before their inflammatory bowel disease is discovered⁵, or that some patients with inflammatory bowel disease clearly have a component of irritable bowel syndrome to account for their symptoms when the inflammatory bowel disease is in remission⁶.

Irritable bowel syndrome and inflammatory bowel disease may be associated with dysbiosis^{7,8} which may account for the abnormal odour emitted from the faeces of patients with both irritable bowel syndrome and inflammatory bowel disease. The volatile

chemicals contributing to faecal odour are mostly products of digestion and fermentation performed by the microbiota and cells shed into the intestine^{9,10}.

Traditionally, volatile chemicals are characterised by gas chromatography - mass spectrometry. Several publications based on gas chromatography - mass spectrometry have shown changes in volatile chemicals found in faeces¹¹, urine¹² and breath¹³ during relapse of inflammatory bowel disease and in faeces of patients with diarrhoea-predominant irritable bowel syndrome¹⁴. These studies give an indication of the potential use of volatile chemicals as biomarkers for inflammatory bowel disease and irritable bowel syndrome, however, the gas chromatography - mass spectrometry technology is not yet suitable for high-throughput applications in clinical practice, which has limited the utility of these observations.

We have designed and built a prototype based on gas chromatography-sensor technology for the point of care analysis of volatile chemical profiles from biological samples. We have reported the preliminary analysis of faecal samples from patients with irritable bowel syndrome and inflammatory bowel disease using the gas chromatography-sensor system and an in-house developed artificial neural network (ANN)¹⁵. However, some important comparisons from the medical point of view were not performed (e.g. active Crohn's disease vs. irritable bowel syndrome and active Crohn's disease vs. inactive Crohn's disease). In addition, the use of ANNs for diagnostic methods has been questioned by regulatory institutions such as the Food and Drug Administration (FDA)^{16,17}.

Here, we report the use of a gas chromatography-sensor-pipeline¹⁸ (a data processing procedure) to analyse faecal samples from patients with irritable bowel syndrome, inflammatory bowel disease and healthy donors. After rigorous validation schemes, the results reported by the gas chromatography-sensor-pipeline indicate a successful discrimination of faecal samples from patients with irritable bowel syndrome, active inflammatory bowel disease, inactive inflammatory bowel disease, active Crohn's disease, active Crohn's disease, active colitis, inactive ulcerative colitis and healthy donors. These results support the development of a point of care device not only for the positive diagnosis of irritable bowel syndrome, but also to assist in the diagnosis of both Crohn's disease and ulcerative colitis.

2 - Methods

2.1 - Patient recruitment

Patients were recruited as described by Shepherd *et al.* 2014¹⁵, although several patients were excluded from the present work as the diagnosis of inflammatory bowel disease was subsequently questioned. In summary, patients attending the gastroenterology clinic at the Bristol Royal Infirmary were invited to participate in this study and to bring a faecal sample to the clinic. Prospective demographic data and faecal samples were obtained from 152 different participants between October 2010 and October 2011.

Irritable bowel syndrome samples include samples from patients with diarrhoea or constipation and patients alternating between diarrhoea and constipation. Most patients had diarrhoea predominant irritable bowel syndrome, however, two patients reported constipation as the predominant symptom. The diagnosis was based on the Rome II criteria¹⁹.

The inflammatory bowel disease samples were collected from patients with active and non-active ulcerative colitis and Crohn's disease. Inflammatory bowel disease was diagnosed by a physician based on endoscopy and histology, or by radiology in the case of small intestinal disease. The activity of the disease in patients with ulcerative colitis was calculated by their colitis simple clinical activity index score²⁰, where a score of 3 or more indicated active UC. Patients with Crohn's disease were assessed using the Harvey Bradshaw index score²¹, where a score of 4 or more indicated active Crohn's disease. Simple clinical activity index has been compared to other tools and found to be "valid, reliable and responsive"²²: it has the advantage over most tools of not requiring an assessment of the mucosa by sigmoidoscopy/colonoscopy. The use of Harvey Bradshaw index is supported by National Institute for Health and Care Excellence (NICE) in the assessment of Crohn's disease patients for anti-TNF therapies: it does not require a diary to be kept by patients for several days, invasive investigations or blood tests. Faecal calprotectin was not measured because it was not a routinely available test in 2010/11. We are not able to perform the test now because the samples collected for this study were disposed of after the sensor work had been completed, in accordance with the Human Tissue Act.

Healthy control samples (Control) (n=41) were collected from partners or healthy relatives of patients visiting the clinic and from healthy patients referred for early endoscopy / colonoscopy due to a family history of upper gastrointestinal or colon cancer; (mean

age 53.6 y, 24 women: 17 men): because of the similarity of their diet and lifestyle to that of patients, partners were recruited where possible as an attempt to reduce bias resulting from such factors. The patients who agreed to participate in the study gave verbal consent to the physician during the clinic appointment as stipulated in the participant information sheet and the ethics approval, as granted by the Wiltshire Research and Ethics Committee (NRES 06/Q2008/6). All patients were on an *ad lib* diet before sample collection in order to maximise recruitment and to give 'real-world' data.

2.2 – Sample processing

All the samples were analysed by the gas chromatography-sensor system in 2012, the device not being available prior to 2012. Faecal samples were processed following the method proposed by Ahmed et al 2013²³. In summary, 1gm aliquots of faecal samples were stored in 18 ml glass headspace vials (Supelco, Sigma Aldrich) within 6h of sample production and frozen at -20°C. In 2012, samples were processed by the system. Previous studies showed no loss of volatile chemicals from faecal samples stored at -20°C²³. Each frozen sample was heated for 10 min at 50°C. After this, 2 cm³ of its headspace were collected and injected into the GC column of the gas chromatography-sensor system¹⁵. Detailed descriptions of the hardware and software¹⁸ are reported elsewhere. In summary, the gas chromatography-sensor system is composed of a gas chromatography column coupled to a metal oxide gas sensor. The sensor is controlled via an electronic circuit monitored by computer software, which records the electrical

resistance of the sensor at 0.5 second intervals during each 40 minute machine run. The resistance profile of each sample generated by the gas chromatography-sensor system was stored in individual text files.

2.3 - Statistical analysis

The gas chromatography-sensor data generated in 2012 were analysed by a new pipeline in 2015/6. A thorough description of the pipeline used here for statistical analysis is described in Aggio et al. 2016¹⁸. In summary, the gas chromatography-sensor characterizes the volatile chemicals present in biological samples. It produces a profile of the sensor resistance vs. time, which describes how the abundances of volatile compounds change with time. Supplementary Figure S1 is a illustrative plot of the average normalised resistance for each of the irritable bowel syndrome, inflammatory bowel disease, and Control (healthy patients) samples (data normalised to be between 0 and 1).

Our in-house-developed pipeline performs chromatogram alignment and data transformation techniques for highlighting volatile chemical patterns specific to different medical conditions. The features or resistance levels that best describe the differences between medical conditions are selected by 2 random forest-based algorithms^{24,25}. Partial least squares (PLS)²⁶ and support vector machine (SVM) with polynomial kernel²⁵ were applied as statistical modelling techniques to classify unknown samples using the derived features. The results reported by the gas chromatography-sensor-pipeline were validated using leave-one-out cross-validation, 10-fold cross-validation repeated 30 times²⁷, 3-fold double cross-validation repeated 30 times with an inner loop of 2-fold

cross-validation repeated 5 times²⁸, and their Monte Carlo variation with random class labels permutation. Principal component analysis (PCA) on the transformed resistance values was also performed. Receiver operating characteristic (ROC) curves were generated based on the double cross-validation results in order to visualise the performance of the gas chromatography-sensor-pipeline. Statistical analyses were performed solely on the resistance profiles processed by the gas chromatography-sensor-pipeline. No other demographic or clinical features were considered for statistical modelling. Confidence intervals (CI) were calculated using bootstrapping. Data analysis was carried out using R software²⁹.

This study is based on data from n = 152 different patient samples comprising data from Controls (n = 41), irritable bowel syndrome (n = 28), and inflammatory bowel disease (n = 83). Pairwise comparisons were performed between these three groups. For detailed comparisons inflammatory bowel disease is further considered as active (n = 33) or inactive (n = 50) and compared with Controls and irritable bowel syndrome. The inflammatory bowel disease data comprises n = 47 ulcerative colitis (active n = 14; inactive n = 33) and n = 36 Crohn's disease (active n = 19; inactive n = 17) and these four subgroups of inflammatory bowel disease are compared with the data from Controls and the irritable bowel syndrome donors. A listing of the comparisons is given in Supplementary Table S1.

We have applied an in-house-developed gas chromatography-sensor-pipeline to analyse 152 faecal samples from patients with irritable bowel syndrome, active inflammatory bowel disease, inactive inflammatory bowel disease, active Crohn's disease, inactive Crohn's disease, active ulcerative colitis, inactive ulcerative colitis and health donors or Control. Table 1 shows the demographics for the patient groups studied with their respective diagnosis, site of disease, Harvey Bradshaw index scores and simple clinical activity index (SCAI) score, when applied, smoking status, diet, medication and routine laboratory data.

Supplementary Table S2 contains a summary of the results reported by the double cross-validation for each comparison performed and Supplementary Table S3 contains the results of their associated ROC analysis. The results reported by the leave-one-out cross-validation, 10-fold cross-validation and Monte Carlo are available as Supplementary Tables S4, S5, S6, S7 and S8. For example, Figures 1 and 2 show the features selected for the comparisons Active- Crohn's disease | irritable bowel syndrome and Active- inflammatory bowel disease | Inactive-inflammatory bowel disease, respectively, in addition to their associated plot of principal components and ROC curves. The results indicate that the platform is able to successfully differentiate most of the conditions studied here, with Active-Crohn's disease | irritable bowel syndrome being an example of near perfect sample classification and Active-inflammatory bowel disease |Inactive-inflammatory bowel disease being an example of a scenario where the platform has difficulty in classifying samples. The prototype device we have built is able to distinguish faecal samples from healthy donors, patients with irritable bowel syndrome and patients with inflammatory bowel disease; the sensitivity and specificity for each is shown in Supplementary Table S2.

The pattern recognition software we have developed is based on wavelet transformation and does not rely on a neural network. In contrast to neural networks, which has been described as a 'black box' approach^{16,17}, the wavelet transformation underpins the technology used to interpret electrocardiogram, a well-known methodology accepted by the scientific community. We have used repeated double cross-validation to validate results. Furthermore, we have undertaken Monte Carlo randomisation to ensure the model is not over-fitted to the data. This is the first time these stringent methods have been used to report faecal volatile compound profiles. These results are supported by Supplementary Figure S1B, which is a plot of average resistance normalised between 0 and 1 over the time 240 seconds to 540 seconds. Supplementary Figure S1B shows distinctive signature differences in average profiles between irritable bowel syndrome, inflammatory bowel disease and Controls over a sustained period.

Making the diagnosis of irritable bowel syndrome, the second most prevalent gastrointestinal disease of westernised populations, is problematic: despite the introduction of the Manning Criteria in 1978³⁰ and the numerous updates of the Rome Criteria many clinicians still feel that irritable bowel syndrome is a diagnosis of exclusion⁴. The introduction of faecal calprotectin has helped to 'rule out' disorders such as inflammatory bowel disease, but still treats irritable bowel syndrome as a diagnosis of exclusion. The

data we have presented is the best method to date for making a positive diagnosis of irritable bowel syndrome based on an investigation. The new pipeline is an improvement on the previously reported neural network, for active inflammatory bowel disease (Crohn's disease and ulcerative colitis) vs irritable bowel syndrome, the pipeline has a mean sensitivity and specificity of 93% and 90%, respectively, the neural network had mean values of 76% and 88%; for irritable bowel syndrome vs controls, the mean accuracy of pipeline and neural network were 91% and 54%, respectively; while for inflammatory bowel disease (combined Crohn's disease and ulcerative colitis) vs controls the mean accuracies were 78% and 79% for the pipeline and neural network, respectively. When Crohn's disease and ulcerative colitis were analysed separately, the pipeline had an accuracy of 89%. This assessment was not undertaken when using the neural network.

The new analysis also compared patients with active Crohn's disease and ulcerative colitis for the first time. The traditional Partial Least Squares (PLS) approach gave a mean accuracy of 94% with an area under the ROC of 99%, SVM gave 96% and 99%, respectively. There are no faecal markers with an ability to distinguish Crohn's disease and ulcerative colitis, although some serology panels show promise³¹. We do not expect faecal volatile compounds to replace standard diagnostic tools such as colonoscopy and MRI or capsular endoscopy, but they could be used to help direct the choice of investigation. Importantly, the technique appears to provide a tool for diagnosing irritable bowel syndrome in a positive way, which will be of reassurance to patients, while saving them from unnecessary tests to rule out other conditions, saving time, money and risk to the patients.

Clinically, the most challenging comparison is between irritable bowel syndrome and active Crohn's disease since both cause abdominal pain and a change in bowel habit. The gas chromatography-sensor pipeline performed well for this comparison with area under the ROC of 91% and 94% using PLS and SVM, respectively, after double cross-validation. The same performance was observed when classifying irritable bowel syndrome and active inflammatory bowel disease, inactive inflammatory bowel disease, inactive Crohn's disease, or inactive ulcerative colitis samples (Supplementary Tables S2 and S3). The assessment of active inflammatory bowel disease / inactive inflammatory bowel disease is rarely useful, as all patients ought to have a clear diagnosis of Crohn's disease or ulcerative colitis; the models were relatively poor and reflect the mixed nature of ulcerative colitis and Crohn's disease patients in the inflammatory bowel disease group (Figure 2). We have chosen to use this figure to emphasise that the profiles for inactive and active inflammatory bowel disease do overlap. More useful were the comparisons of inactive/active ulcerative colitis or Crohn's disease, here the models were better, especially for ulcerative colitis (Supplementary Figure S2) in which all but one sample from patients with active ulcerative colitis lay to the right of the vertical line on the principal component plot.

The clinical assessment of disease activity in ulcerative colitis is more accurate than that of Crohn's disease, because the colon is more readily assessed than the small bowel and the Crohn's disease activity index, a commonly used scoring index in clinical trials, is very subjective. The samples were collected in 2010-11. Faecal calprotectin testing was not routinely available at the research centre, which means we had no robust measure of disease activity. In addition, faecal calprotectin has its limitations in the

assessment of small bowel Crohn's disease. Consequently, we chose two patient-friendly but reliable clinical tools, the simple clinical activity index and Harvey Bradshaw index. Future work will need to assess the performance compared with a robust gold-standard such as colonoscopy for ulcerative colitis, or colonoscopy with MRI for Crohn's disease.

The results reported here were validated using the following different validation schemes: leave-one-out cross-validation, 10-fold cross-validation, double cross-validation and Monte Carlo randomisation. Extant statistical literature holds all of these methods and approaches in good standing. Among them, the double cross-validation is certainly the most stringent method. This stringent approach performed least well when comparing inactive/active Crohn's disease, or ulcerative colitis. This is not unexpected for two reasons in addition to the sample size; (1) the change in the volatile compounds are a continuous variable that was compared to an arbitrary cut point in a second continuous variable (Harvey Bradshaw index or simple clinical activity index) – comparing upper and lower quartiles may have been more discriminating, but the data set was too small for this; (2) patients may have had other reasons for their symptoms (such as bile salt diarrhoea, bacteria overgrowth or irritable bowel syndrome) which meant the clinical scores over-estimated disease activity.

The Monte Carlo technique was applied in order to check for potential over-fitting of the developed classification models. The Monte Carlo method was applied as described for the validation methods tested (i.e. leave-one-out cross-validation, cross-validation and double cross-validation), however, in this case, sample labels were randomly permuted before model construction.

This procedure simulates what would have happened if samples were to be classified simply by chance. The results (Supplementary Tables S4, S6 and S8) suggest the data were not over-fitted.

We have developed a gas chromatography-sensor pipeline for the diagnosis and assessment of inflammatory bowel disease and irritable bowel syndrome. The separation of active disease groups is excellent. The potential to use the pipeline to determine disease activity will require more work. Although the sample sizes were too small to provide separate validation sets, we have used stringent statistical tools to double cross validate our models. We are planning further large cohort studies with gold-standard assessments of disease activity and validation sets. If confirmed these findings could mean that irritable bowel syndrome can be diagnosed positively and offers the potential to develop new tools to diagnose and assess inflammatory bowel disease and distinguish ulcerative colitis and Crohn's disease.

Acknowledgements

We are grateful to patients and colleagues for providing us with the clinical samples. We would like to thank Miss Sophie Shepherd for providing technical support by aliguoting and loading samples.

Figure caption:

Figure 1. Volatile compounds were extracted from faecal samples and analysed using an in-house-developed gas chromatography – sensor device. The profiles of volatile chemicals were further analysed using an in-house-developed computer algorithm that identifies resistance patterns associated with specific medical conditions. This figure shows the selected features used to differentiate 19 samples from patients with active Crohn's disease (Act-CD) and 28 patients with irritable bowel syndrome (IBS), and a two-component principal component plot based on the selected features.

Figure 2. Volatile compounds were extracted from faecal samples and analysed using an in-house-developed gas chromatography – sensor device. The profiles of volatile compounds were further analysed using an in-house-developed computer algorithm that identifies resistance patterns associated with specific medical conditions. This figure shows the Receiver operating characteristic (ROC) curve and the principal component analysis based on the selected features used to differentiate 33 samples from patients with active inflammatory bowel disease (Act-IBD) and 50 patients with inactive inflammatory bowel disease (Inact-IBD). The ROC analysis was based on repeated double cross-validation using support vector machine polynomial and partial least squares.

References

1 - Spiller R, Aziz Q, Creed F. Guidelines on the irritable bowel syndrome: Mechanisms and practical management (vol 56, pg 1770, 2007). *Gut* 2008;**57**:1743-.

2 - Longstreth GF, Thompson WG, Chey WD, et al. Functional bowel disorders. Gastroenterology 2006;130:1480-91.

3 - Drossman DA, Camilleri M, Mayer EA, Whitehead WE. AGA technical review on irritable bowel syndrome. *Gastroenterology* 2002;**123**:2108-31.

4 - Spiegel BMR, Farid M, Esrailian E, Talley J, Chang L. Is irritable bowel syndrome a diagnosis of exclusion?: A survey of primary care providers, gastroenterologists, and IBS experts. *American Journal of Gastroenterology* 2010;**105**:848-58.

5 - Porter CK, Cash BD, Pimentel M, Akinseye A, Riddle MS. Risk of inflammatory bowel disease following a diagnosis of irritable bowel syndrome. *BMC Gastroenterology* 2012;**12**.

6 - Halpin SJ, Ford AC. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: Systematic review and meta-analysis. *American Journal of Gastroenterology* 2012;**107**:1474-82.

7 - Casen C, Vebo HC, Sekelja M, *et al.* Deviations in human gut microbiota: A novel diagnostic test for determining dysbiosis in patients with IBS or IBD. *Alimentary Pharmacology & Therapeutics* 2015;**42**:71-83.

8 - Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease. *Microbial Ecology in Health and Disease* 2015;**26**:26191-.

9 - Zlatkis A, Brazell RS, Poole CF. The role of organic volatile profiles in clinical-diagnosis. *Clinical Chemistry* 1981;27:789-97.

10 - Probert CSJ, Ahmed I, Khalid T, *et al.* Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases. *Journal of Gastrointestinal and Liver Diseases* 2009;**18**:337-43.

11 - Ahmed I, Greenwood R, Costello B, Ratcliffe N, Probert CS. Investigation of faecal volatile organic metabolites as novel diagnostic biomarkers in inflammatory bowel disease. *Alimentary Pharmacology & Therapeutics* 2016;**43**:596-611.

12 - Cauchi M, Fowler DP, Walton C, *et al.* Comparison of GC-MS, HPLC-MS and SIFT-MS in conjunction with multivariate classification for the diagnosis of Crohn's disease in urine. *Analytical Methods* 2015;**7**:8379-85.

13 - Bodelier AGL, Smolinska A, Baranska A, *et al.* Volatile organic compounds in exhaled air as novel marker for disease activity in Crohn's disease: A metabolomic approach. *Inflammatory Bowel Diseases* 2015;**21**:1776-85.

14 - Ahmed I, Greenwood R, Costello Bdl, Ratcliffe NM, Probert CS. An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome. *Plos One*. 2013;**8**(3).

15 - Shepherd SF, McGuire ND, de Lacy Costello BPJ, *et al.* The use of a gas chromatograph coupled to a metal oxide sensor for rapid assessment of stool samples from irritable bowel syndrome and inflammatory bowel disease patients. *Journal of Breath Research* 2014;**8**.

16 - Rodvold DM. Validation and regulation of medical neural networks. *Molecular Urology* 2001;5:141-5.

17 - Lisboa PJ, Taktak AFG. The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* 2006;**19**:408-15.

18 - Aggio RBM, de Lacy Costello B, White P, *et al.* The use of a gas chromatography-sensor system combined with advanced statistical methods, towards the diagnosis of urological malignancies. *Journal of Breath Research* 2016;**10**:017106-.

19 - Drossman DA. The functional gastrointestinal disorders and the Rome II process. Gut. 1999;45:1-5.

20 - Walmsley RS, Ayres RCS, Pounder RE, Allan RN. A simple clinical colitis activity index. Gut 1998;43:29-32.

21 - Harvey RF, Bradshaw JM. A simple index of Crohn's disease activity. *Lancet* 1980;1:514.

22 - Turner D, Seow CH, Greenberg GR, et al. A systematic prospective comparison of noninvasive disease activity indices in ulcerative colitis. *Clinical Gastroenterology and Hepatology* 2009;7:1081-8

23 - Ahmed I, Smith S, Probert CS. Volatile organic compounds as diagnostic faecal biomarkers in inflammatory bowel disease method development. *Gut* 2009;**58**:A63-A.

24 - Kursa MB, Rudnicki WR. Feature selection with the boruta package. Journal of Statistical Software 2010;36:1-13.

25 - Kuhn M. Caret: Classification and regression training. R package version 6.0-30 edn., 2014.

26 - Geladi P, Kowalski BR. Partial least-squares regression - a tutorial. Analytica Chimica Acta 1986;185:1-17.

27 - Delen D. Analysis of cancer data: A data mining approach. *Expert Systems* 2009;**26**:100-12.

28 - Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. Journal of Chemometrics 2009;23:160-71.

29 - Team RC. R: A language and environment for statistical computing.: R Foundation for Statistical Computing., 2015.

30 - Manning AP, Thompson WG, Heaton KW, Morris AF. Towards positive diagnosis of irritable bowel. *British Medical Journal* 1978;**2**:653-4.

31 - Plevy S, Silverberg MS, Lockton S, *et al.* Combined serological, genetic, and inflammatory markers differentiate non-IBD, Crohn's disease, and ulcerative colitis patients. *Inflammatory Bowel Diseases* 2013;**19**:1139-48.

Tables

Table 1. Summary of demographic of patients with inflammatory bowel disease and irritable bowel syndrome

		Crohn's disease		Ulcerative colitis			IBS	Controls
		Inactive CD N=17	Active CD N=19		Inactive UC N=35	Active UC N=14	N=28	N=41
Mean Age (range)		47.9 (19-67)	44.0 (23-63)		60.8 (22-90)	56.8 (34-81)	43.1 (18-73)	53.6 (17-94)
Gender F		7	8		13	6	19	24
Ν	1	10	11		20	8	9	17
	L1	5	6	E1	4	4	NA	NA
Cite of disease	L2	8	12	E2	14	6		
Site of disease	L3	2	1	E3	15	4		
	L4	2						
Median disease activity (range)	0 (0-3)	8 (4-22)		0 (0-2)	6 (3-9)	NA	NA
Median CRP (n, range)		3.0(6, 0.9-12)	9.5(12, 2-40)		2 (14, 1-25)	5 (10, 0.9-6)	3.5 (8, 1-9)	1 (10, 1-5)
Median WCC (n, range)		5.6(7, 3.5-8.2)	7.4 (12, 6.3-17.8)		7(18, 3.8-10.2)	6.8(10, 4.5-13.6)	7.4 (10, 3.5-9.7)	6.2 (10, 5.4-12.7)
Median PV (n)		1.7 (4)	1.8 (10)		1.7 (12)	1.7 (10)	1.6 (8)	NA
Current medication Steroids		2	8			3		
Azathiopine / Mercaptopurine		2	3		5	2		
Methotrexate		1	2		1			
Infliximab		2	1					
Adalimumab		0	1					

5-ASA	3	2	25	9		
Salazopyrine	2	0				
Iron						4
Movicol					2	
Smoker: Yes/No/ Ex	6/1/11	1/16	2/28/5	1/9/4	NA	NA
Diet: Mixed / Veg/Polymeric	17/0/0	17/1/1	33/2/0	11/0/0	NA	NA

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy	4
		(such as sensitivity, specificity, predictive values, or AUC)	
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	4
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	5,6
	4	Study objectives and hypotheses	7
METHODS			
Study design	5	Whether data collection was planned before the index test and reference standard	7,8
		were performed (prospective study) or after (retrospective study)	
Participants	6	Eligibility criteria	7
	7	On what basis potentially eligible participants were identified	7
		(such as symptoms, results from previous tests, inclusion in registry)	
	8	Where and when potentially eligible participants were identified (setting, location and dates)	7
	9	Whether participants formed a consecutive, random or convenience series	7
Test methods	10a	Index test, in sufficient detail to allow replication	8,9,10
	10b	Reference standard, in sufficient detail to allow replication	8,9,10

	11	Rationale for choosing the reference standard (if alternatives exist)	8,9,10
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8,9,10
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8,9,10
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8,9,10
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8,9,10
Analysis	14	Methods for estimating or comparing measures of diagnostic accuracy	8,9,10
	15	How indeterminate index test or reference standard results were handled	8,9,10
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	NA
	18	Intended sample size and how it was determined	NA
RESULTS			
Participants	19	Flow of participants, using a diagram	NA
	20	Baseline demographic and clinical characteristics of participants	10
	21a	Distribution of severity of disease in those with the target condition	10
	21b	Distribution of alternative diagnoses in those without the target condition	10
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	11

	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	11
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	11,12,13,14,15
	27	Implications for practice, including the intended use and clinical role of the index test	11,12,13,14,15
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	7,8
	30	Sources of funding and other support; role of funders	1