
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/147192/

Deposited on: 11 September 2017

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk
Luminally-expressed Gastrointestinal Biomarkers

G. Cummins¹, D. E. Yung², B. F. Cox³, A. Koulaouzidis², M. P. Y. Desmulliez¹, S. Cochran⁴

1. Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom

2. The Royal Infirmary of Edinburgh, Endoscopy Unit, Edinburgh, EH16 4SA, United Kingdom

3. School of Medicine, University of Dundee, Dundee, DD1 9SY, United Kingdom

4. Medical and Industrial Ultrasonics, School of Engineering, University of Glasgow, G12 8QQ, United Kingdom

Corresponding author details:

Email: G.Cummins@hw.ac.uk

Journal:

Expert review of gastroenterology and hepatology
1. Abstract

Introduction

A biomarker is a measurable indicator of normal biologic processes, pathogenic processes or pharmacological responses. The identification of a useful biomarker is challenging, with several hurdles to overcome before clinical adoption. This review gives a general overview of a range of biomarkers associated with inflammatory bowel disease or colorectal cancer along the gastrointestinal tract.

Areas covered

These markers include those that are already clinically accepted, such as inflammatory markers such as faecal calprotectin, S100A12 (Calgranulin C), Fatty Acid Binding Proteins (FABP), malignancy markers such as Faecal Occult Blood, Mucins, Stool DNA, Faecal microRNA (miRNA), other markers such as Faecal Elastase, Faecal alpha-1-antitrypsin, Alpha2-macroglobulin and possible future markers such as microbiota, volatile organic compounds and pH.

Expert commentary

There are currently a few biomarkers that have been sufficiently validated for routine clinical use at present such as FC. However, many of these biomarkers continue to be limited in sensitivity and specificity for various GI diseases. Emerging biomarkers have the potential to improve diagnosis and monitoring but further study is required to determine efficacy and validate clinical utility.
2. **Keywords**

biomarker  
gastrointestinal disease  
faecal calprotectin  
lactoferrin  
phagocyte derived protein,  
faecal occult blood testing,  
mucins,  
fatty acid binding proteins,  
microbiome,  
volutilome

3. **Introduction**

A biomarker is defined as an objectively measurable characteristic that is objectively measured and evaluated as an indicator of normal biologic or pathogenic processes, pathogenic processes or pharmacological responses to a specified therapeutic intervention [1]–[3]. Ideally it should be measured from easily obtained bodily fluids and waste such as plasma, serum, urine and faecal matter for improved patient acceptance; if not feasible then invasive techniques such as tissue biopsy are acceptable. An ideal biomarker would be sensitive to a specific condition. A useful biomarker has the potential to improve diagnosis, discriminate ill patients according to disease, reduce healthcare costs, expedite drug development and monitor therapeutic efficacy.

In this review, a general overview is given of established and potentially new biomarkers found within the intestinal lumen, such as proteins, enzymes, microbes and their metabolic products (summarised in Table I) are assessed in terms of their usefulness for the diagnosis of conditions such as colorectal cancer (CRC) and inflammatory bowel disease (IBD). Combined with advances in point of care diagnostics and minimally invasive therapies, these biomarkers may be expected to improved patient outcomes in the future.
4. Luminally expressed gastrointestinal biomarkers

4.1. Markers of inflammation

4.1.1. Phagocyte-derived proteins as markers of inflammation

Markers derived from phagocytes are useful as indicators of inflammation, with those expressed only by phagocytes being more specific. Markers which are also inducible in epithelial cells are more sensitive, but their levels can be raised by other non-inflammatory but still physiologically stressful conditions, such as lactose intolerance. Neutrophil activation markers have been found to be involved in the pathogenesis of inflammatory bowel disease (IBD), therefore their levels often show good correlation with disease activity. However, cytokines such as the interleukins and tumour necrosis factor-alpha (TNFα) tend to be short-lived and are unstable molecules, and methods for their detection are unreliable [4], [5].

Lactoferrin, polymorphonuclear (PMN)-elastase, myeloperoxidase (MPO) and human neutrophil lipocalin are markers of neutrophil degranulation, which can be detected in stool, and have therefore been proposed as markers of gastrointestinal (GI) tract inflammation. Lactoferrin, an iron binding glycoprotein, is considered the most accurate of these, but is also found in epithelial cells and consequently has limited specificity. In a large meta-analysis [6], lactoferrin was found to have a sensitivity of 82% and specificity of 79% for IBD, compared to a sensitivity of 49% and specificity of 92% seen in C-reactive protein (CRP). How lactoferrin is involved in inflammatory pathways is not well-established [7]–[9]. PMN-elastase has been used to determine the severity of acute pancreatitis, with the advantage of being elevated earlier than conventional markers such as CRP. Markers of eosinophil degranulation, mainly eosinophilic cationic protein and eosinophilic protein X, have been found to have some use in intestinal hypersensitivity and eosinophilic inflammation [4].

A major drawback of using phagocyte-derived proteins as markers of inflammation is their lack of specificity to the GI tract, as they are expressed during any inflammatory process. For example, neutrophil MPO has found use in assessing other inflammatory conditions including lung inflammation in smokers [10], [11].
4.1.2. **Faecal Calprotectin**

The S100 proteins are a family of phagocyte-specific damage-associated molecular pattern molecules (DAMPs). They encompass a family of over 20 known calcium-binding proteins with tissue-specific expression patterns. S100A8, S100A9 and S100A12 are specifically linked to innate immune functions by their expression in phagocytes [4].

The S100A8/S100A9 complex is better known as faecal calprotectin (FC). FC is a calcium and zinc binding heterodimer protein with a regulatory role in the inflammatory process, and is released during cell activation and cell death. It constitutes about 60% of soluble proteins in human neutrophilic cytosol and is also found in monocytes, macrophages and ileal tissue eosinophils. It is excreted into the gut lumen as a response to inflammatory stimuli and is therefore six times more concentrated in faeces than in plasma and other bodily fluids [12].

FC is currently in wide clinical use as a marker of IBD activity, both in screening patients with suspected IBD and monitoring disease activity in those with an established diagnosis. In daily clinical practice, FC has proven useful due to the ease of collection of stool samples; the complex is stable at room temperature, resistant to degradation and homogeneously distributed in stool [8], [9], [12]–[14]. The current literature supports the use of FC as a screening and monitoring biomarker for both main forms of IBD – ulcerative colitis (UC), characterised by mucosal colonic inflammation, and Crohn’s disease (CD), which can affect all parts of the gastrointestinal tract with transmural inflammation. The cut-off point for “normal” FC is usually taken to be 50 µg/g of stool, as specified by the test manufacturers; however cut-off values vary widely depending on clinical situation and even these values are still a matter of debate and study. Conventionally, cut-off values of 50-100 µg/g have been reported in the literature, depending on individually calculated receiver operative curves in different studies. The management of patients with “borderline” FC levels (50-150 µg/g) remains unclear as many such patients are not eventually diagnosed with IBD [15]. A summary of the
various reported FC levels corresponding to different clinical scenarios in IBD is given in Table II.

The use of FC as a screening test in suspected IBD has been shown to reduce the number of endoscopies with negative results – in a meta-analysis by van Rheenen et al [6], the pooled sensitivity of FC for IBD in adults was 93%, with pooled specificity 96%, and the authors estimated that screening with FC would reduce the number of endoscopies required in these adults by 67%. Another meta-analysis [16] found that using a cut-off of FC of 50 µg/g gave lower sensitivity and specificity (89% and 81%, respectively) compared to FC of 100 µg/g (sensitivity 98%, specificity 91%). Gisbert and McNicholl [17] reported that FC had a higher diagnostic accuracy for CD (sensitivity 83%, specificity 85%) than UC (sensitivity 72%, specificity 74%). Interestingly, these studies and other subsequent ones [18] have found a lower specificity of FC for IBD in children of about 70% although sensitivities were comparable.

FC is also useful in monitoring disease activity in patients with established IBD. Studies have shown that FC correlates well with endoscopic disease activity measured using a number of different indices including the Rachmilewitz index [19], the modified Baron index in UC [20], and Crohn’s Disease Index of Severity (CDIS) [21], [22]. The normalisation of FC has good correlation with mucosal healing which is an important prognostic indicator in IBD [23]–[27]. By now, several studies have demonstrated that patients with both UC and CD who were in remission following medical treatment also had normalisation of FC levels to below 50 µg/g, whereas non-responders continued to have elevated FC [28]. Conversely, rising FC concentrations in patients with clinically quiescent disease have been found to predict clinical relapse with up to 90% sensitivity and 82% specificity. A large meta-analysis by Mosli et al in 2015 [6] showed overall sensitivity of 88% and specificity 73% for endoscopically active IBD. Sensitivities were similar between UC (88%) and CD (87%), but the specificity of FC was lower in CD (67%) than UC (79%); this is in line with previous studies [12], [29]. Furthermore, mildly elevated levels of FC in unaffected relations of individuals with IBD show its sensitivity to even subclinical GI tract inflammation [30], [31].
In general, the optimal levels or cut-offs of FC to use vary widely across different clinical scenarios and studies. Furthermore, there has been some emerging evidence that FC levels may display intra-individual variability from day to day, or even within the same day, observed in individuals with active IBD [32], [33], quiescent IBD [34] and also normal individuals [35]. Therefore it is suggested that more attention needs to be paid to timing of sample collection and the practicalities of sample storage [32], and also that multiple or serial FC measurements [33] are obtained. Optimal cut-offs may also differ by disease (UC, CD), brand of assay used, distribution of inflammation, and age of patient [36]. Clinical guidance from various organisations including the British Society of Gastroenterology [37] and the European Crohn’s and Colitis Organisation [38] does not contain a clear-cut definition of “normal” FC and recommendations seem to be guided by the trend of an individual patient’s FC levels, i.e. increasing or normalising.

Despite its ease of use and wide adoption, FC is not specific to IBD. Other causes of GI tract inflammation can result in elevated FC, including use of non-steroidal anti-inflammatory drugs (NSAIDs) [39]–[41] and malignancy [42]–[44]. Another limitation of FC is its inability to localise pathology [43], [44]. It is currently thought that FC is more reflective of disease activity in UC than in CD, as discussed above, and in colonic CD compared to small bowel CD [31]. However, a recent meta-analysis has shown that elevated FC in patients with normal ileocolonoscopy was associated with active small bowel inflammation [45].

4.1.3. S100A12 (Calgranulin C)

S100A12 (calgranulin C) has recently emerged and may be more specific for IBD than FC. Studies have shown that, unlike FC and lactoferrin, it is elevated in IBD but not in conditions such as gastroenteritis. This is thought to be because S100A12 is expressed only by granulocytes and acts independently of FC in calcium-dependent signalling, hence also being thought to be involved in the pathogenesis of IBD [14], [46], [47]. Elevated S100A12 >10 mg/kg
has been reported to have sensitivity of 96% and specificity of 92% for IBD in children, and sensitivity of 86% and specificity of 96% in adults [28].

4.1.4. Other markers
Potential biomarkers are constantly being discovered [48]–[50]. Recent studies on neopterin demonstrated its ability to identify patients with IBD with active mucosal lesions and could potentially be used to assess the severity their mucosal damage [51]. It was shown that FC and neopterin concentrations correlated closer with endoscopic scores for UC ($r = 0.75$ and $r = 0.72$, respectively; $p < 0.0001$ for both) than in CD ($r = 0.53$ and $r = 0.47$, respectively; $p < 0.0001$ for both). Neopterin has a similar overall accuracy to FC in predicting endoscopic activity in IBD patients when cut-offs of 250 μg/g for FC and 200 pmol/g for Neopterin were used.

4.2. Fatty acid binding proteins (FABP)
Fatty acid binding proteins (FABPs) are low molecular weight cytosolic proteins found in tissues involved in the uptake and consumption of fatty acids. Three forms have been found to be expressed by enterocytes: intestinal FABP (I-FABP), liver FABP (L-FABP) and ileal-bile acid binding protein (I-BABP). As they are found in the tips of villi, the presence of FABPs in the circulation or urine has been considered as a marker of early enterocyte damage [52]. I-FABP levels have been shown to increase with acute intestinal ischaemia and inflammation, implying correlation with ischaemia-reperfusion injury [52]. For example, a rise in urinary I-FABP following cardiopulmonary bypass was found to be associated with increased risk of postoperative complications [53]. A significant inverse association has also been demonstrated between I-FABP and I-BABP and mean arterial pressure, thought to be due to intestinal mucosal cell injury secondary to systemic hypotension and consequent hypoperfusion [52]. Elevated plasma I-FABP is also associated with poorer outcomes in patients with abdominal sepsis [54].
The usefulness of these FABPs as GI biomarkers is due to their specificity for gut pathology and early elevation following enterocyte damage, allowing a rapid response in time-critical conditions such as sepsis. Their excretion in urine is also helpful, especially in the paediatric setting.

4.3. Markers of malignancy

4.3.1. Faecal Occult Blood Testing

Faecal occult blood testing (FOBT) detects traces of blood in stool. The earliest form of FOBT, guaiac (gFOBT), is nonspecific for human blood as it is based on the detection of pseudoperoxidase activity [55], [56]. Consequently it has relatively low sensitivity and specificity; for example, false positives could be obtained due to recent consumption of red meat [57]. Therefore, faecal immunochemical tests (FITs), which specifically and quantitatively detect human blood through antibody reaction with human globin, are now in wider use. FITs are also thought to be more specific for distal GI blood loss such as in CRC. Another form of FIT is haemoglobin-haptoglobin (Hb/Hpt) complex testing. Haptoglobin forms a stable soluble complex with haemoglobin which is detected via an immunoradiometric assay [58]. A large systematic review on the various available forms of FOBT found a wide range of reported sensitivities, from 6.2-83.3% for gFOBT and 5.4-62.6% for FITs [59]. Specificity performed better, from 65.0-99.0% for gFOBT and 89.4-98.5% for FITs.

FOBT is widely used worldwide to screen for adenomas and CRC. The use of FOBT has been shown to increase early detection of CRC and improve survival [57], [60], [61]. A Cochrane review in 2008 [61] concluded that FOBT reduced the relative risk of mortality from CRC by 16%, with risk reduction increased to 25% when results were adjusted for screening attendance. A more recent Cochrane review [62] compared FOBT to flexible sigmoidoscopy for the detection of CRC; both modes of screening reduced the relative risk of death from CRC (0.72 in flexible sigmoidoscopy vs 0.86 with FOBT). Although flexible sigmoidoscopy
understandably performed better, FOBT offers advantages as it is cheap and easily implemented and therefore suitable for screening large numbers of people efficiently.

There is little published data on the value of FOBT in predicting small bowel pathology, although it is thought to be less useful [63]. 40-60% of subjects with positive FOBT have no lesions found in the colon or rectum with colonoscopy, raising the possibility of small bowel bleeding [63], [64].

4.3.2. Mucins: MUC1, MUC2, MUC5AC, small-intestinal mucin antigen, CDX2, villin

Mucins are high molecular weight glycoproteins expressed in various types of epithelium. Specific mucins such as MUC1, MUC2 and MUC5AC have been used as markers for small bowel and CRC, as the development and progression of neoplasia are associated with aberrant mucin expression [65]–[68]. Similarly, small-intestinal mucin antigen (SIMA) is an oncofetal glycoprotein present only in normal small bowel in adults; its abnormal expression is associated with adenocarcinomas of the GI tract [65], [69]. Therefore, such markers can be useful in distinguishing the types of detected neoplasia in patients and help to guide further management.

Villin is an actin-binding protein involved in brush border maintenance that is specific for epithelial neoplasms in the GI tract. It has proven useful in distinguishing between GI tract cancer and ovarian, bladder and prostate cancers. Another sensitive marker for GI adenomas and especially CRC is CDX2, a homeobox nuclear transcription factor [65]. However there remains a dearth of information about the accuracy parameters and therefore utility of these markers and they are currently not in routine use.

4.4. Faecal or Stool DNA (sDNA)

During the evolution of normal colonocytes into CRC, an accumulation of genetic mutations and epigenetic alterations occurs within aberrant colonocytes [70]. These cells are continuously exfoliated into the lumen and passed in faeces [71].
Genetic variations are represented by various mutations [72] such as KRAS and APC. Additional changes are represented by epigenetic aberrant hyper- or hypo-methylation [73]. Aside from their role in the genesis and propagation of CRC, these modified molecules have a demonstrated role as biomarkers. This is due in part to their presence in plasma/serum [74], urine [75] and stool [76]. Furthermore it has been noted that the mutated genetic material remains stable in the stool [77] and can be further stabilised ex vivo using buffers [78]. Amplification and means of detection employ variations of polymerase chain reaction (PCR) [79], [80].

As a screening modality, sDNA has demonstrated variable sensitivity and specificity with performance linked to the markers analysed and in what combination they are analysed. A meta-analysis by Zhai et al [81], including 53 studies, calculated a pooled sensitivity of 48% and specificity of 97% for single gene testing. Multigene testing experienced an improved pooled sensitivity of 77.8% with a slightly lower specificity of 92.7%. The authors speculate that the improved multigene sensitivity is due to evaluating sDNA more broadly while simultaneously maintaining a precise mutation search. Compared to FIT, multigene sDNA demonstrated superior sensitivity but lower specificity when performed in an asymptomatic, average risk individual [82]. In a large cross-sectional study (n = 9,989), Imperiale and colleagues compared FIT against multitarget sDNA (MT-sDNA) for aberrant KRAS, NDRG4, BMP3 and β-actin for DNA quantity reference. Post stool analysis, participants underwent colonoscopy as the reference standard. Sensitivity for CRC detection was 92.3% for MT-sDNA and 73.8% for FIT (p = 0.002). Detection of advanced precancerous lesions measuring ≥ 1 cm, including advanced adenomas and sessile serrated polyps, resulted in sensitivities of 42.4% and 23.8% for MT-sDNA and FIT, respectively (p < 0.001). Specificity of sDNA was 86.6% and 94.9% for FIT (p < 0.001). Approval by the FDA of MT-sDNA testing was granted in August, 2014 [83].

In addition to its utility in screening for CRC, sDNA may have a potential role in prognosis and as a predictive biomarker [84]. This opinion is echoed by Okugawa et al [17] with particular
attention directed towards CpG island methylator phenotype (CIMP) as a promising prognostic indicator. This is due to CIMP positive tumours correlating with a negative prognosis as indicated by Phipps et al [85]. Furthermore CIMP positive status was employed as a predictor of 5-FU response [86] though this has since been refuted [87]. However a meta-analysis by Juo et al [88] was inconclusive due to lack of standardisation and recommended further follow-up with a RCT.

Comparative-effectiveness studies are now needed to clarify the role of stool DNA testing with respect to programmatic screening with other test options. Only through a better understanding of other key factors, such as the screening interval, adherence, cost, and diagnostic evaluation of positive results, can we determine the appropriate place for stool DNA testing on the screening menu [89].

4.5. Faecal microRNA (miRNA)

MicroRNA (miRNA) is a short single stranded string of non-coding RNA [90]. This family is 18-25 nucleotides in length and plays a role in the regulation of oncogenes and tumour suppressors [91]. Altered miRNA expression, in the form of up or down regulation, has been linked to a number of cancers including CRC [92], as well as to IBD [93]. The small hairpin structure of miRNA renders it resistant to RNAse. It thus demonstrates a high degree of stability in a number of media including blood and faeces [94], [95] and can be recovered from fixed and frozen tissue [96]. Detection and quantitation is by PCR and microarray analysis with other means being investigated [97].

Several groups have run pilot studies aimed at investigating miRNA profiles and biomarker potentials in human stool samples [95], [98], [99]. Wu at al compared miRNA-21 and -92a levels in pre- and post-surgery CRC patients, individuals with polyps and healthy controls [100]. It was first noted that stool based miRNA-92a detection was robust and reproducible over a 72 hour time span. Significantly higher miRNA levels were noted in CRC patients than in healthy controls (p < 1.01 and p < 0.0001 for miRNA-21 and -92a respectively). Furthermore, miRNA-92a was significantly elevated in the patients with polyps group versus
the healthy controls ($p < 0.0001$). Interestingly, tumour removal resulted in a reduction of both miRNA-21 and -92a ($p < 0.01$) and a decrease in miRNA-92a for advanced adenoma excision ($p < 0.05$). Further investigation compared miRNA-135b amongst four different groups, which included patients with CRC, adenomas, IBD and healthy controls [100]. The use of an IBD group was to restrict markers to individuals with neoplastic disease. From an array analysis, miRNA-31 and 135b were identified as the most unregulated. MiRNA-135b was found to be significantly higher in patients with CRC ($p < 0.0001$) and adenomas ($p < 0.0001$) than in the IBD and control groups. Furthermore, an increasing trend was noted along the adenoma to carcinoma spectrum ($p < 0.0001$). Additional results showed that miRNA was indifferent to proximal or distal location of the lesion and target stool levels dropped post excision. The sensitivity and specificity of miRNA-135b for CRC were 78% and 68%, respectively. Sensitivity for adenomas was 65% and further improved to 73% for advanced adenomas.

Additional evaluation of miRNA and comparison to FIT was done by Koga et al [98]. Tests were carried out on confirmed CRC patients and healthy controls. Sensitivity and specificity for miRNA-106a alone were 34.2% and 97.2% respectively versus significant ($p = 0.001$) FIT results of 60.7% and 98.1%, respectively. However when combined, sensitivity was elevated to 70.9% with a slight drop in specificity to 96.3%. Further investigation by Ahmed et al discovered the over and under-expression of a number of miRNA targets [99]. These results illustrate the feasibility of miRNA as a non-invasive biomarker for colon-related neoplasia screening, due to the feasibility of collecting stool as well as the change in miRNA expression levels along the adenoma-carcinoma spectrum, correlating with TNM staging

### 4.6. Markers of other pathology

#### 4.6.1. Faecal Elastase

Faecal elastase-1 is an enzyme specific to the human pancreas. It is currently used as a test for exocrine pancreatic insufficiency as it is not degraded during intestinal transport and is therefore testable in a stool sample [101], [102]. Its faecal concentration has been found to correlate with levels of pancreatic enzyme secretion but sensitivity and specificity have been
shown to vary. Faecal elastase performs better in moderate to severe pancreatic insufficiency but has been shown to be superior to other tests of pancreatic function [102]–[104].

4.6.2. Faecal alpha-1-antitrypsin
Faecal alpha-1-antitrypsin is a marker of GI plasma loss or protein losing enteropathy, with the advantage that it is less invasive than the use of radiolabelled macromolecules. Furthermore, it shows good correlation with protein loss and is a relatively inexpensive test [5]. It was previously used as a marker of CD but its use has declined with the introduction of FC and other more specific markers [105], [106].

4.6.3. Alpha2-macroglobulin
Alpha2-macroglobulin is a protease which has been used as a marker for protein losing enteropathy and pancreatitis. It has been suggested that increased protease activity and consumption in certain inflammatory conditions such as acute pancreatitis, sepsis and IBD could cause its levels to fall [5]. However, it is highly non-specific and there is little conclusive evidence to support its routine use.

4.7. Microbiota
There are approximately $100 \times 10^{12}$ microorganisms within the GI tract, an order of magnitude greater than the number of somatic cells within the body. These microorganisms primarily reside within the large intestine and include bacteria, yeasts, single eukaryotes and more. This section will primarily discuss the role of bacteria on GI health. Several factors aid bacterial growth within the large intestine, such as the slow transit time of the colon, neutral pH and low bile salt concentration [107]. Bacteria are located within two sections of the large intestine [107]–[109]. As the colon is devoid of oxygen, most of this bacterial population is anaerobic. The bacteria within the large intestine provide several useful functions such as protection of epithelial cells, aid the immune system, stimulation of intestinal angiogenesis, and fermentation of nondigestible dietary fibre, cellulose, resistant starches, gums and pectins, leading to the production of short-chain fatty acids (SCFA) which are an energy source for colonocytes [89],[90].
Whilst reflective of the community in the colonic lumen, the microbiome in the luminal section may not reflect the composition of the epithelial and cryptal community. The microbiome in the mucosa is a dense community of bacteria that adheres to the surface. That can withstand the hydrodynamic shear forces present [111]. Adherent resident bacteria may play a role in the development of IBD and CRC.

The Human Microbiome Project and other research show that there is a high degree of variation in the bacterial populations between individuals [2] [112] and studies have shown that factors such as age, diet, ethnicity and environment may contribute to this variation. Despite the dissimilarities, many of these bacteria that carry out common metabolic activities are similar between different individuals [113]. Commonly found genera of bacteria within the adult large intestine include *Bifidobacterium, Eubacterium, Lactobacillus, Bacteroides, Clostridium, Escherichia, Streptococcus and Ruminococcus*. *Bacteroides* is one of the most abundant within the GI tract, with *Eubacterium, Bifidobacterium, Peptostreptococcus, and Ruminococcus* also widespread [107],[114].

An increasing number of studies have demonstrated that there is a relationship between changes in the gut microbial population and the development of diseases such as IBD [96]–[101] and CRC [107], [110], [111], [122]–[137]. A summary of some of these studies is shown in Table III. Many of the trials shown in this table use a relatively small number of samples, which limits the wider applicability of the results. Additionally, several studies have identified changes in different bacterial populations to be linked to the same GI disease. One contributing factor may be that many of the studies shown in Table III were conducted using stool samples as a proxy for mucosal samples due to ease of collection. However it is unknown to what extent the faecal and mucosal microbiomes differ, which may have influenced the results of these studies [138]. Besides the use of faecal biota as opposed to mucosal biota, other factors that may limit the widespread applicability of these studies may include lack of information regarding microbiota patchiness, and heterogeneity between different anatomical
niches along the colon, difference due to ages and genders, and other factors, limit the usefulness and [2].

It is unclear, whether the changes in gut microbiota contribute to disease pathogenesis or whether they occur due to local inflammation [139]. The bacterial driver-passenger model proposed by Tjalsma could explain the influence of microbiota on the development of GI disease [140]. This model suggests that certain populations of driver bacteria with pro-carcinogenic features damage intestinal epithelium DNA, leading to tumorigenesis, which in turn alters the intestinal environment leading to a decline in homeostasis due to overgrowth of opportunistic passenger bacteria.

Study of the gut microbiome is still in its infancy and the lack of a holistic understanding of this complex environment and its role in gut pathology presents a major barrier to the identification of bacteria or combinations of bacteria as potentially reliable, specific and sensitive biomarkers for CRC and IBD. Currently, identification of the microbial constituents can be performed through use of molecular fingerprinting methods and sequence analysis of cloned microbial ribosomal DNA due to difficulties in cultivation. These methods are expensive and take time, which currently impedes the recognition of widely accepted biomarkers for clinical usage. The development of inexpensive, minimally invasive, rapid and reliable technologies for direct sampling and analysis of the microbiome will aid the design and operation of clinical trials necessary in identifying these biomarkers and utilising them to aid diagnosis in future.

4.8. Volatile Organic Compounds

Direct intra-luminal measurement of microbiota population and demographics may not be possible with current technology. However, the competition between different populations of bacteria for resources such as substrates or hydrogen can lead to deviations in the concentrations of metabolic by-products such acetic, propionic and butyric acids, CO₂, H₂, CH₄, NH₃, H₂S and volatile fatty acids [141]. These changes in metabolic activity may be reflected in changes in the constituent gases of the patients of breath. Diagnosis by breath analysis is an area of much interest as measurements are non-invasive and potentially near
real time, making it ideal for screening or point of care diagnosis. However, this is not trivial as human breath is a complex mixture of gases containing at least 3000 different compounds [142], which includes more than 800 volatile organic compounds (VOCs). Changes in the relative concentration of the VOCs may correlate with pathological or physiological changes [143].

Several studies have proposed various VOC biomarkers for CRC or IBD [144]–[157] with the results of some of these studies shown in Table IV. However, the sample size for these studies has been small, which may be a factor in different biomarkers being identified for the same pathologies. Further study is needed amongst a larger population to identify useful, widely applicable biomarkers for the detection of GI disease.

The rapid development and expansion of capsule endoscopy functionality through increased use of microengineering has enabled integration of various sensing modalities, such as gas sensors into these diagnostic tools [158], [159]. While, the prospect of direct or indirect sampling and subsequent analysis of the VOCs produced within the GI tract for clinical diagnosis is attractive, there is a need for further large, randomised and blinded studies before reliable, sensitive and specific VOC markers can be definitively identified for the diagnosis of GI pathology in clinical applications.

4.9. pH

Changes in luminal pH along the GI tract is an established subject of interest. The pH values of a typical, healthy GI tract is generally agreed to be as shown in Table V, showing an increase from the duodenum to the terminal ileum, a decrease in the caecum and a slow rise along the colon to the rectum. The values can change based on time since time since ingestion, diet, age and other factors [160],[161]. Changing pH along the GI tract occurs for several reasons, such as the absorption of acetate, propionate, butyrate and other short chain fatty acids, fermentation, and production of alkaline metabolites [114].
It has been proposed that UC and to some extent other IBDs such as CD can cause a decrease in pH due to mucosal inflammation within the colon [158]-[160], but this is subject to debate [162]. Drawing conclusions is difficult due to the small size of the studies, variations in dietary intake, extent and severity of the colitis, and possible signal loss or drift due to the use of endoscopic capsules. Several studies have shown that CRC is not linked with any significant change in pH [163]-[165] despite an earlier hypothesis [166] that acidification of the bowel may play a role in the aetiology of CRC.

5. Expert Commentary

Endoluminal biomarkers such as FC and lactoferrin have changed the management of GI disease, allowing detection at an earlier stage or improved diagnostic certainty. The increased understanding of GI disease pathogenesis and progression due to the emergence of omics in recent years has led to the identification of thousands of potential biomarkers. However, there remain significant limitations to their use. These limitations include the low sensitivity and/or specificity of tests like FOBT and current lack of comprehensive understanding of the complexities of the gut microbiome. Other barriers to clinical adoption include clinician education in the conduct and interpretation of tests, and the relevant costs. Therefore, only a few biomarkers have been sufficiently validated for routine clinical use at present. Nevertheless, emerging biomarkers have the potential to improve clinical practice, for example accurate VOC biomarkers could improve the rate of screening of CRC in the wider population due to the non-invasive and rapid nature of breath analysis, various stool biomarkers could offer non-invasive testing with quicker results than serum markers. The study of these biomarkers and development of testing methods could also further increase understanding of GI disease mechanisms and therefore improve management, patient care and outcomes.

6. Five Year View

The transfer of biomarkers from benchtop to clinical practice is not trivial. To become clinically approved, an effective biomarker must be validated using hundreds of samples, be specific -
able to correctly identify a high proportion of true negative rates; sensitive - able to correctly identify a high proportion of true positives and also be reproducible - able to achieve the same level of detection under similar conditions multiple times. Validation of biomarkers in independent cohorts from several different institutions will help this process. To date, no marker has been found with 100% sensitivity and specificity for any GI disease. However, several solutions have been proposed to help biomarkers progress from the initial preclinical exploratory stage to final control studies. These include improving the assay used for detection, effective use of biomarker combinations and identification of the subpopulations in which the biomarker is most effective for use as a stratified diagnostic tool. This will be helpful for verifying potential VOC and microbiota biomarkers that have so far been identified in small studies, prone to bias due to various factors such as diet, ethnicity and age. Over the next five years, the increasing adoption of inexpensive, point of care diagnostic tools in labs will enable quicker screening and characterisation of effective biomarkers or combinations of biomarkers for various GI diseases, thereby speeding up the route to clinical adoption.

7. Key Issues
- There are currently well-established biomarkers such as FC. However, biomarkers continue to be limited in sensitivity and specificity for various GI diseases.
- Several biomarkers show potential but further study is needed to determine their efficacy, validate their use and overcome barriers to clinical adoption.
- New and emerging GI biomarkers could improve disease diagnosis and monitoring through better sensitivity, specificity and/or ease and speed of measurement.
- A greater understanding of the role of the microbiome in GI disease is important for future biomarker identification and more effective use of existing but not yet well-established biomarkers.
- The ideal biomarker is easy to obtain from patients (e.g. stool, urine or breath samples as opposed to serum) and has an inexpensive test, which is easy to conduct and interpret, with excellent sensitivity and specificity.
Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under the grant number EP/K034537/1 and entitled Sonopill.
<table>
<thead>
<tr>
<th>Type</th>
<th>Biomarker</th>
<th>Possible Cause</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflammation markers</td>
<td>Phagocyte derived protein</td>
<td>Inflammatory bowel disease</td>
<td>Good sensitivity to IBD</td>
<td>Lack of specificity to inflammation of the GI tract</td>
<td>[5], [8]–[10], [167], [168][6]</td>
</tr>
<tr>
<td></td>
<td>Faecal Calprotectin</td>
<td>Inflammation</td>
<td>Well established, relatively inexpensive</td>
<td>Moderate specificity for IBD</td>
<td>[8], [11]–[29][6] [15] [16] [17] [18] [19] [20] [21] [22] [23]–[27][28]</td>
</tr>
<tr>
<td></td>
<td>S100A12 (Calgranulin C)</td>
<td>Inflammatory bowel disease</td>
<td>Initial studies show greater specificity than faecal calprotectin for IBDA</td>
<td>Larger studies required</td>
<td>[14], [46], [47]</td>
</tr>
<tr>
<td></td>
<td>Fatty acid binding proteins</td>
<td>Inflammation associated with early stage enterocyte damage</td>
<td>Good specificity for gut pathology</td>
<td>Larger studies required</td>
<td>[52]–[54]</td>
</tr>
<tr>
<td>Malignancy markers</td>
<td>Faecal Occult Blood Testing</td>
<td>Distal GI blood loss</td>
<td>Popular and inexpensive method of screening</td>
<td>Low sensitivity and specificity</td>
<td>[33]–[37], [38], [39], [58] [59] [61] [62]</td>
</tr>
</tbody>
</table>

Table I: Summary of biomarkers
<table>
<thead>
<tr>
<th>Faecal DNA</th>
<th>KRAS</th>
<th>Colorectal cancer</th>
<th>Initial studies show moderate sensitivity and good specificity.</th>
<th>Larger studies required.</th>
<th>[45]–[55], [56], [57], [59]–[63]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CIMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecal microRNA</td>
<td>miRNA-21</td>
<td>Colorectal cancer</td>
<td>Initial studies show good sensitivity and specificity.</td>
<td>Larger studies required</td>
<td>[90]–[99] [100].</td>
</tr>
<tr>
<td></td>
<td>miRNA-92a</td>
<td></td>
<td>Trend towards reduction in cost of analytical equipment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>miRNA-106a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>miRNA-135b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enzymes</td>
<td>Faecal elastase-1</td>
<td>Pancreatic function</td>
<td>Good correlation with pancreatic enzyme secretion.</td>
<td>Varied sensitivity and specificity.</td>
<td>[101]–[104]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Demonstrated superior performance to other tests of pancreatic function.</td>
<td>Limited to moderate and severe pancreatic insufficiency</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Faecal alpha-1-antitrypsin</td>
<td>GI plasma loss or protein losing enteropathy</td>
<td>Good correlation with protein loss.</td>
<td>Less specific than faecal calprotectin and other markers that have superceded it in terms of usage</td>
<td>[105], [106][5]</td>
</tr>
<tr>
<td>Other</td>
<td>Alpha2-macroglobulin</td>
<td>Protein losing enteropathy and pancreatic, sepsis and inflammatory bowel disease</td>
<td>Good correlation with acute pancreatitis, sepsis and IBD</td>
<td>Non specific</td>
<td>Little conclusive evidence to support routine use</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>-------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Microbiota</td>
<td>Colorectal cancer, inflammatory bowel disease</td>
<td>See Table III for further information</td>
<td>Knowledge of the role of microbiota on GI pathology still in its infancy</td>
<td>Further studies required to identify reliable microbiota biomarkers</td>
<td>See Table III for further information</td>
</tr>
<tr>
<td>Volatile Organic Compounds</td>
<td>Potential for minimally invasive detection of GI disease without need to handle stool or urine samples. See Table IV for further information</td>
<td>Knowledge of the volatilome still basic, requiring further studies to link changes in gas concentration to pathology. More accurate and inexpensive methods of gas detection also required. See Table IV for further information</td>
<td>[144]–[157]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>Inflammatory bowel disease</td>
<td>Further study required due to confounding factors such as age, diet and other factors.</td>
<td>[162]–[166], [176]–[180]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table II: Summary of use of faecal calprotectin in IBD

<table>
<thead>
<tr>
<th>Clinical scenario</th>
<th>FC levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnosis</strong></td>
<td>&gt;50 usually used to distinguish between IBD and functional disorders; however 50-150 is generally considered a grey area</td>
</tr>
<tr>
<td></td>
<td>&gt;100 has better specificity but sensitivity varies between studies; is considered “strongly positive”</td>
</tr>
<tr>
<td><strong>Active vs inactive IBD</strong></td>
<td>&gt;100: endoscopically active CD</td>
</tr>
<tr>
<td></td>
<td>&gt;250: large ulcers seen on endoscopy</td>
</tr>
<tr>
<td></td>
<td>&lt;250: mucosal healing on endoscopy/ endoscopic remission</td>
</tr>
<tr>
<td></td>
<td>In general, normalisation of FC correlates with mucosal healing.</td>
</tr>
<tr>
<td><strong>Treatment response</strong></td>
<td>Rise in FC &gt;100 predicts clinical relapse</td>
</tr>
<tr>
<td></td>
<td>Normalisation of FC or at least &lt;75% reduction from baseline can be used as surrogate marker for mucosal healing</td>
</tr>
<tr>
<td></td>
<td>Normalisation of FC to below 100 may predict clinical remission</td>
</tr>
<tr>
<td><strong>Postoperative recurrence in CD</strong></td>
<td>&gt;100-200 (depending on study)</td>
</tr>
<tr>
<td></td>
<td>In general, a rise in FC is considered significant</td>
</tr>
<tr>
<td></td>
<td>FC expected to normalise within 2 months after uncomplicated ileocaecal resection; persistently raised FC could reflect ongoing inflammation</td>
</tr>
<tr>
<td><strong>Relapse</strong></td>
<td>&gt;250 measured 3 months after diagnosis: high relapse risk in UC</td>
</tr>
<tr>
<td></td>
<td>&gt;500: thought to distinguish high relapse risk from low relapse risk (paediatric study)</td>
</tr>
</tbody>
</table>

### Table III: Summary of experiments demonstrating relationship between changes in microbiome and presence of disease

<table>
<thead>
<tr>
<th>Disease</th>
<th>Bacteria</th>
<th>Sample</th>
<th>Subjects</th>
<th>Clinical Outcome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crohn’s Disease</td>
<td>Faecalibacterium prausnitzii, <em>Escherichia coli</em></td>
<td>Biopsy samples</td>
<td>DNA analysis of biopsy collected from 5 locations between ileum and rectum from 6 discordant monozygotic twin pairs against 4 concordant pairs</td>
<td>Patients with ileal Crohn’s disease had a reduced population of Faecalibacterium prausnitzii (<em>p</em>&lt;0.001) and an abundance of <em>Escherichia coli</em> (<em>p</em>&lt;0.03) compared to healthy discordant twins and those with colon localised Crohn’s</td>
<td>[172]</td>
</tr>
<tr>
<td>Crohn’s Disease</td>
<td><em>Escherichia coli</em></td>
<td>Biopsy samples</td>
<td>28 patients were studied: 13 with Crohn’s disease involving the ileum, 8 with Crohn’s ileal mucosa found to have a population of <em>Escherichia coli</em> (<em>P</em>&lt;0.001), but relatively depleted in a subset of Clostridiales (<em>P</em>&lt;0.05)</td>
<td></td>
<td>[181]</td>
</tr>
<tr>
<td>Disease</td>
<td>Bacterial Species</td>
<td>Sample Type</td>
<td>Sample Details</td>
<td>Observed Effect</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Crohn's Disease</td>
<td><em>Butyricicoccus pullicaecorum</em></td>
<td>Stool samples</td>
<td>51 patients with Crohn’s disease, 25 with Ulcerative Colitis versus 88 healthy controls</td>
<td>The average number of <em>Butyricicoccus</em> bacteria was significantly (p&lt;0.0001) lower in the stools of IBD group versus healthy controls. A significantly lower level of <em>Butyricicoccus</em> species was observed in the faecal microbiota of patients with active CD compared with CD in remission (p&lt;0.0188)</td>
<td></td>
</tr>
<tr>
<td>Crohn's Disease</td>
<td><em>Bdellovibrio bacteriovorus</em></td>
<td>Biopsy samples</td>
<td>9 patients with Crohn’s, 6 patients with Ulcerative Colitis versus 8 healthy controls</td>
<td>Statistically higher population of <em>B. bacteriovorus</em> in ileum, colon, and rectum of control biopsies with respect to CD (p&lt;0.0001). No difference was found among overall UC and control samples p=0.6760.</td>
<td></td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td><em>Faecalibacterium prausnitzii</em></td>
<td>Stool samples</td>
<td>127 patients with Ulcerative Colitis (39 active, 88 in remission) versus 87 healthy controls</td>
<td>Real-time PCR analysis revealed a lower abundance of <em>Roseburia hominis</em> and <em>Faecalibacterium prausnitzii</em> (p&lt;0.0001) in UC patients compared to controls</td>
<td></td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td><em>Roseburia hominis</em></td>
<td>Biopsy samples</td>
<td>8 patients (4 with colon cancer and 4 with rectal cancer). Study was performed with 16 tissue samples taken during colonoscopy, 8 taken from healthy tissue and 8 taken from cancerous tissue.</td>
<td>50% of Chinese CRC patients, we found a significant increase of Roseburia (p = 0.017), and a concurrent decrease of both Microbacterium (p = 0.009) and Anoxybacillus (p = 0.009) in tumor tissue</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td><em>Fusobacteria Eubacteriaceae Clostridiales Family XI. Incertae sedis Staphylococcaceae Bacteroides Campylobacteraceae Porphyromonadaceae Enterococcaceae</em></td>
<td>Stool samples</td>
<td>19 patients with CRC versus healthy controls</td>
<td>Increased population of <em>Fusobacteria</em> (p&lt;0.01), <em>Eubacteriaceae</em> (p=0.037), <em>Clostridiales Family XI. Incertae sedis</em> (p = 0.004), <em>Staphylococcaceae</em> (p=0.011), <em>Bacteroides</em> (p=0.046) compared to healthy control group. Reduced population of <em>Campylobacteraceae</em> (p=0.014) and <em>Porphyromonadaceae</em> (p=0.001) families compared to healthy control group. No significant change in <em>Enterococcaceae</em> (p=0.062).</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td><em>Bacteroides/Prevotella</em></td>
<td>Stool samples</td>
<td>Stool samples taken from 60 patients with confirmed colorectal cancer after colonoscopy versus healthy controls</td>
<td>Elevated population of <em>Bacteroides/Prevotella</em> in CRC patients compared to healthy controls (p = 0.009)</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td><strong>Bacteroides</strong></td>
<td><strong>Roseburia</strong></td>
<td><strong>Alistipes</strong></td>
<td><strong>Eubacterium</strong></td>
<td><strong>Parasutterella</strong></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td><strong>Atopobium</strong></td>
<td><strong>Clostridia</strong></td>
<td><strong>Fusobacterium</strong></td>
<td><strong>Porphyromonas</strong></td>
<td>Stool samples</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td><strong>Acidaminobacter</strong></td>
<td><strong>Phascolaractobacterium</strong></td>
<td><strong>Citrobacter farmer</strong></td>
<td><strong>Akkermanasia muciphilia</strong></td>
<td><strong>Bacteroides finegoldii</strong></td>
</tr>
<tr>
<td>Disease</td>
<td>Volatile Organic Compounds</td>
<td>Method of Analysis</td>
<td>Subjects</td>
<td>Clinical Outcome</td>
<td>Reference</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------</td>
<td>--------------------</td>
<td>----------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Methane</td>
<td>-</td>
<td>45 patients with colorectal cancer versus healthy controls</td>
<td>No significant difference between colorectal cancer and control groups</td>
<td>[144]</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Methane</td>
<td>-</td>
<td>59 patients with unresected colorectal cancer versus healthy controls</td>
<td>No significant difference between colorectal cancer and control groups</td>
<td>[145]</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Methane</td>
<td>GC</td>
<td>55 patients with unresected colorectal cancer versus healthy controls</td>
<td>No significant difference observed between patients with colorectal cancer and healthy controls</td>
<td>[146]</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Methane</td>
<td>GC</td>
<td>47 patients with unresected colorectal cancer, 36 patients with resected colorectal cancer, 7 with nonresectable cancer, 29 with non-malignant diseases of colon</td>
<td>The majority of patients (91.4%) with unresected colorectal cancer produced more methane than healthy controls (p&lt;0.001) and patients with benign diseases of the colon (p&lt;0.001)</td>
<td>[147]</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Ethanol, Acetone, Ethyl acetate, 4-methyl octane</td>
<td>GC-MS</td>
<td>65 patients with colorectal cancer, 22 with adenoma versus health controls</td>
<td>Increased Acetone and Ethyl acetate in CRC patients (p = 0.010, p = 0.005 respectively) versus healthy controls. Reduced Ethanol and 4-methyl octane in CRC patients (p&lt;0.001, p=0.004 respectively) versus healthy controls.</td>
<td>[148]</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Non-anal 4-Methyl-2-pentanone Decanal 2-Methylbutane 1,2-Pentadiene 2-Methylpentane 3-Methylpentane Methylcyclopentane Cyclohexane Methylcyclohexane 1,3-Dimethylbenzene 4-Methyloctane 1,4-Dimethylbenzene 4-methylundecane trimethyldecane</td>
<td>GC-MS</td>
<td>37 patients with colorectal cancer versus controls</td>
<td>Elevated concentrations detected in CRC patients compared to control</td>
<td>[149]</td>
</tr>
<tr>
<td>Disease</td>
<td>Compounds</td>
<td>Method</td>
<td>Patients</td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>1,10-(1-butenylidene)bis benzene, 1,3-dimethyl benzene, 1-iodo nonane, [(1,1-dimethylethyl)thio] acetic acid, 4-(4-propylcyclohexyl)-40-cyanq(1,10-biphenyl)-4-yl ester benzoic acid, 2-amino-5-isopropyl-8-methyl-1-azulenecarbonitrile</td>
<td>GC-MS</td>
<td>26</td>
<td>Elevated concentrations detected in CRC patients compared to controls</td>
<td></td>
</tr>
<tr>
<td>Inflammatory Bowel Disease</td>
<td>Pentane</td>
<td>GC</td>
<td>33</td>
<td>Correlation with disease activity</td>
<td></td>
</tr>
<tr>
<td>Inflammatory bowel disease a</td>
<td>1-Octene, 1-Nonene, 1-Decene, z3- Methylhexane, (E)2-Nonene, Hydrogen sulphide</td>
<td>SIFT-MS</td>
<td>62</td>
<td>Elevated readings of 1-octene, 1-decene, 3-methylhexane and reduced readings of 1-Nonene, (E)-2-Nonene, hydrogen sulphide in IBD patients compared to control. P&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Inflammatory Bowel Disease</td>
<td>2-propanol, acrylonitrile, carbon disulfide, dimethylsulfide, ethanol, isoprene, triethylamine</td>
<td>SIFT-MS</td>
<td>24</td>
<td>Changes in concentrations of gases observed compared to healthy controls (p&lt;0.001 for all gases). There was no significant difference in any VOC levels between CD and UC</td>
<td></td>
</tr>
<tr>
<td>Irritable Bowel Syndrome</td>
<td>Hydrogen Sulphide</td>
<td>MS</td>
<td>27</td>
<td>Patients confirmed to have small bowel intestinal overgrowth (SIBO) in irritable bowel syndrome showed higher concentrations of H2S compared to those that were negative SIBO (p&lt;0.001)</td>
<td></td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>Ethane, Pentane</td>
<td>GC</td>
<td>17</td>
<td>Significantly higher compared to controls (p&lt;0.013). Positive correlation of ethane with endoscopic score, symptom score, disease activity and chemiluminescence in rectal tissue. Pentane levels did not correlate with any of the clinical measurements</td>
<td></td>
</tr>
<tr>
<td>Ulcerative colitis and Crohn's disease</td>
<td>Ethane, Propane, Butane, Isoprene</td>
<td>GC</td>
<td>10</td>
<td>Significant difference between elevated patients with IBD and controls were found for ethane, propane, and pentane (p ranges between &lt;0.05 to &lt;0.001), but no</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- **Irritable Bowel Syndrome**
  - Hydrogen Sulphide

**References:**
- [150]
- [151]
- [152]
- [153]
- [154]
- [155]
- [156]
<table>
<thead>
<tr>
<th>Ulcerative Colitis, Colorectal cancer, Crohn’s disease</th>
<th>Methane</th>
<th>GC</th>
<th>Decreased methane concentration on breath of UC and CD patients versus healthy controls (p&lt;0.001, p&lt;0.001). Increased methane concentration on breath of CC patients versus healthy controls (p&lt;0.005)</th>
<th>[157]</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 patients with unresected colorectal cancer, 40 patients with ulcerative colitis and 40 patients with Crohn’s disease versus healthy controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organ</td>
<td>pH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stomach</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Intestine</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caecum</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References:


[21] T. Lobatón, A. López-García, F. Rodríguez-Moranta, A. Ruiz, L. Rodríguez, and J. Guardiola, “A new rapid test for fecal calprotectin predicts endoscopic remission and


[112] A. B. Shreiner, J. Y. Kao, and V. B. Young, “The gut microbiome in health and in

A. Buck, M. C. Rivera, N. U. Sheth, S. R. Canon, P. S. G. Chain, C.-C. C. Lo, M.
Scholz, N. C. Kyrpides, K. Liolios, V. M. Markowitz, K. Mavromatis, I. Pagani, I.-M. A.
M. A. Chen, K. Chu, K. Palaniappan, M. A. Cutting, H. A. Hamilton, E. L. Harris, R.
Dwayne Lunsford, P. McInnes, C. C. Davis, T. Z. Desantis, F. E. Dewhirst, J. Izard, K.
P. Lemon, E. Deych, P. S. La Rosa, W. D. Shannon, W. Michael Dunne, M. A.
Forney, J. Friedman, C. S. Smillie, N. Garcia, A. Gonzalez, D. Knights, S. Kinder
Pop, D. D. Sommer, C. M. Lewis, P. Spicer, T. Madden, P. J. Mannon, A. L. McGuire,
S. M. Patel, M. Podar, T. A. Vishivetskaya, K. S. Pollard, T. J. Sharpton, R. M. Truty,
M. Rho, Y. Ye, R. Rhodes, K. P. Riehle, P. Sankar, P. D. Schloss, A. M. Schubert, T.
M. Schmidt, G. A. Simone, J. D. Sobel, T. J. Treangen, S. Yooseph, L. Zoloth, S.
Fulton, M. G. Giglio, K. Hallsworth-Pepin, E. A. Lobos, R. Madupu, V. Magrini, J. C.
Martin, M. Mitreva, D. M. Muzny, E. J. Sodergren, J. Versalovic, A. M. Wollam, K. C.
Worley, J. R. Wortman, S. K. Young, Q. Zeng, K. M. Aagaard, O. O. Abolude, E.
Allen-Vercoe, E. J. Alm, L. Alvarado, G. L. Andersen, S. Anderson, E. Appelbaum, H.
Bhonagiri, M. Bihan, M. J. Blaser, T. Bloom, V. Bonazzi, J. Paul Brooks, G. A. Buck,
C. J. Buhay, D. A. Busam, J. L. Campbell, S. R. Canon, B. L. Cantarel, P. S. G.
Chain, I.-M. A. M. A. Chen, L. Chen, S. Chhibba, K. Chu, D. M. Ciulla, J. C. Clemente,
S. W. Clifton, S. Conlan, J. Crabtree, M. A. Cutting, N. J. Davidovics, C. C. Davis, T.
Z. Desantis, C. Deal, K. D. Deleaunaty, F. E. Dewhirst, E. Deych, Y. Ding, D. J.
Dooling, S. P. Dugan, W. Michael Dunne, A. Scott Durkin, R. C. Edgar, R. L. Erlich, C.
L. J. Forney, L. Foster, V. Di Francesco, J. Friedman, D. C. Friedrich, C. C. Fronick, L.


