# An investigation of volatile organic compounds as biomarkers for gastrointestinal neoplasia

Thesis submitted in accordance with the requirements of the University of Liverpool for an MD by Dr Ashley Bond

**Abstract**

**An investigation of volatile organic compounds as biomarkers for gastrointestinal neoplasia.**

Thesis submitted in accordance with the requirements of the University of Liverpool for an MD by Dr Ashley Bond.

Early identification of malignancy has the ability to improve long term morbidity and mortality. This is certainly the case for colorectal cancer and hepatocellular carcinoma (HCC), two of the commonest gastrointestinal malignancies. Both diseases are subject to screening programmes. The UK national Bowel Cancer Screening Programme (BCSP) has been shown to significantly improve 5 year cancer survival by identifying colorectal cancer at an earlier stage, but also by identifying advanced, pre-malignant adenomatous disease and removing it. Screening of cirrhotic patients at risk of HCC is a bit more controversial but is recommended by national and international bodies, with some evidence to suggest an improved survival from cancer associated death. Biomarkers are currently employed in the diagnosis and patients selection process for these screening programmes, in particular alpha fetaprotein (AFP) for HCC. The sensitivity and specificity, and thus the valid application of these biomarkers has been brought in question, in the case of alpha feta protein, leading to its removal from screening protocols. Volatile organic compounds have been proposed as biomarkers for various disease processes, including gastrointestinal malignancies. They may therefore have an application in disease screening and/or monitoring.

The work presented here explores volatile organic compounds (VOCs) emitted from stool and urine, in order to detect disease specific differences that may be utilised as biomarkers for colorectal cancer and hepatocellular carcinoma. It also explores the driver-passenger model of colorectal cancer and biological plausibility via the detection of volatile organic compounds emitted from cultures of *Fusobacterium nucleatum* and *Campylobacter showae*. Finally, it assesses the utility of the stool based tM2-PK assay as a marker of colorectal neoplasia in a novel secondary care cohort.

Solid phase micro-extraction of headspace gas followed by gas chromatography mass spectrometry was used to isolate and identify candidate volatile organic compounds. Statistical analysis, including logicistic regression modelling and 10 fold cross validation, were applied to assess biomarker utility.

Analysis of VOCs emitted from stool was able to differentiate those with higher risk neoplastic disease with the greatest confidence, including established colorectal cancer. When comparing

those with cancer to no neoplasia isopropyl alcohol was significantly more abundant in the colorectal cancer samples (p=<0.0001, q=0.004), producing an AUROC curve of 0.76. When isopropyl alcohol is combined with butanoic acid, 3-methyl, the AUROC was 0.82, sensitivity 87.9% (95% CI 0.87-0.99) and specificity 84.6% (95% CI 0.65-1.0). Further logisitic regression analysis of VOC presence identified a three VOC panel (isopropyl alcohol, 2-Hexanone and butanoic acid,3-methyl-,ethyl ester) with an AUROC of 0.86: a person being 6 times more likely to have cancer if all 3 VOCs were present in their stool(p=<0.0001). A number of the VOCs identified as important in those with colorectal neoplasia were also identified in the assessment of *Fusobacterium nucleatum* and *Campylobacter showae,* namely butanoic acid based compounds and isopropyl alcohol. VOCs emitted from urine failed to demonstrate any candidate biomarkers for colorectal neoplasia.

With regards VOCs emitted from urine as biomarkers for HCC, AUROC comparing all those with and without HCC was 0.76 (Sensitivity 65% [95% CI 0.61-0.69] Specificity 74% [95% CI 0.69-0.78]). When assessing treatment naive HCC patients, 3 compounds were found to have significantly different abundance (p=<0.01), when combined and modelled these VOCs demonstrated a superior AUROC of 0.81 (Sensitivity 77% [95% CI 0.71-0.83], Specificity 75% [95% CI 0.71-0.79]). Of this treatment naive group patients defined by Barcelona Clinic Liver Cancer (BCLC) staging as having early disease and therefore potentially curative, demonstrated an AUROC of 0.82.

VOC emitted from stool and urine show a clear ability to act as biomarkers for the diagnosis of colorectal cancer and hepatocellular carcinoma respectively.

N.B.

All images not produced by the author have been included with the permission of their copyright owners.

**Acknowledgements**

I would like to thank Prof Probert for giving me the opportunity to work with him over the last two years in my role as a clinical research fellow. He has provided a great deal of academic and professional support allowing me to develop professionally and personally whilst completing my MD. I was able to conduct the research independently whilst always feeling support and guided. I must also thank Prof Pritchard for his help during his time as my supervisor over the two years. I would also like to thank Dr Michael Burkitt for his assistance with various matters and general advice in negotiating the realm of academia over the last two years. Thank you also to Dr Rosemary Greenwood and Dr Georgina Hold for their specific assistances during the undertaking of this research.

I must also acknowledge the team which I joined 2 years ago, namely Raphael Aggio, Sophie Reade and Arno Mayor who were on hand to assist me during the early days of my work and teach me the details of SPME headspace extraction and GCMS.

Finally I would like to say thank you to my wife Clare and our two beautiful children Noah and Emma for their unwavering support.

**List of abbreviations**

AFB- Aflatoxin B1

AFP- Alpha-fetoprotein

AFU- Alpha-I-fucosidase

ALD- alcoholic liver disease

AMDIS- Automated Mass Spectral Deconvolution and Identification System

APC- Adenomatous polyposis coli

AUROC- Area under receiver operating characteristic

BCFA- Branch chain fatty acid

BCLC- Barcelona Clinic Liver Cancer

BCSP- Bowel Cancer Screening Programme

Bft- *Bacteroides fragilis*

BM- Biomarker

BMI- Body mass index

CA19-9- Carbohydrate antigen

CAR-PDMS- Carboxen®/Polydimethylsiloxane

CDT Cytolethal distending toxin

CEA- Carcinoembryonic antigen

CI- confidence interval

CIBH- Change in bowel habit

CIMP- CpG island methylator phenotype

COX 2- Cycloxygenase

CRC- colorectal cancer

CRMP 2- Collapsin response mediator protein 2

CT- Computed topograhpy

CTS- Circadian timing system

CV- Coefficient of variation

DCP- Des-gamma-carboxy prothrombin

DNA- Deoxyribonucleic acid

EASL- European Association for the Study of Liver Disease

EGF- Epidermal growth factor

EGFR- Epidermal growth factor receptor

ELISA- Enzyme-linked immunosorbent assay

EPIC- European prospective investigation into cancer and nutrition

EU- European Union

FAIMS- Field asymmetric ion mobility spectrometer

FDR- False discovery rate

FH- Family history

FIT- Faecal immunological test

FN- False negative

FOBt- Faecal occult blood test

FP- False positive

GCMS- Gas chromatography mass spectrometry

gFOBt- guaiac-based faecal occult blood test

GI- Gastrointestinal

HBV- Hepatitis B virus

HCC- Hepatocellular carcinoma

HCV- Hepatitis C virus

HGF- Hepatocyte growth factor

HS- Headspace

HSD-Tukey's test -Honest significance difference Tukey's test

IARC- International agency for research on cancer

IBD- Inflammatory bowel disease

IBS- Irritable bowel syndrome

IDA- Iron deficiency anaemia

IL- Interleukin

iNOS- Inducible nitric oxide synthase

KRAS- Kirsten rat sarcoma

LOH- Loss of genetic heterozygosity

m/z- Mass-to-charge ratio

M2 PK- Pyruvate kinase isoenzyme type M2

MELD- Model for end stage liver disease

MLL4- Mixed-lineage leukaemia 4

MRI- Magnetic resonance imaging

MSI- Microsatellite instability

NAFLD- Non-alcoholic fatty liver disease

NASH- Non-alcoholic steatohepatitis

NHS- National Health Service

NIST- National Institute of Standards and Technology

NF-kB- Nuclear factor kB

NOS- Nitrogen reactive species

PCA- Principle component analysis

PCR- Polymerase chain reaction

PLS-DA- Partial least square discriminant analysis

RCT- Randomised control trial

RFA- Radio frequency ablation

RNA Ribonucleic acid

ROC- Receiver operating characteristic

ROS- Reactive oxygen species

RR- Relative risk

RT- Retention time

SCCA- Squamous cellular carcinoma antigen

SCFA- Short chain fatty acid

SD- Standard deviation

SEM- Standard error of mean

SOP- Standard operating procedure

SPME- Solid phase micro extraction

TACE- Trans-arterial chemoembolisation

TERT- Telomerase reverse transcriptase

TGF-Beta- Transforming growth factor beta

TIMP 1- Tissue inhibitor of metalloproteinase type 1

TN- True negative

TNF- Tumour necrosis factor

TP- True positive

UK- United Kingdom

USA- United States of America

VEGF- Vascular endothelial growth factor

VOC- Volatile organic compound

WHO- World Health Organisation

## Contents

**Introduction**

## List of Tables

## Table of Figures

# Chapter 1

## Introduction

## An investigation of biomarkers for gastrointestinal neoplasia

### 1.1 Colorectal cancer: epidemiology and statistics

Colorectal cancer is a major cause of mortality and morbidity in Europe and North America, with an estimated European incidence of 43.5 per 100,000 and a mortality of 19.5 per 100,000 in 2012[1]. The incidence of colorectal cancer has increased by 6% in the last decade: it is now the third most common malignancy in the UK[2]. The UK lifetime risk is 1 in 15 for men and 1 in 19 for women[3]. Colorectal cancer accounts for approximately 15,000 deaths per annum in the UK, with a 50-55% 5-year mortality rate[2]. It is associated with a significant financial burden on the UK National Health Service (NHS): the mean annual cost of treating a patient with rectal cancer is estimated to be £12,000, while that for a patient with non-rectal colon cancer is approximately £8,800[4]. Early diagnosis is a key aim in managing colorectal cancer and population based screening has been shown to reduce long term mortality from colorectal cancer when implemented[5].

### 1.2 Acquired risk factors

Many studies have reported associations between dietary and lifestyle factors and the risk of developing colorectal cancer. Others have gone further, with attempts to discover and describe potential, underlying biological processes and mechanisms that could influence genetic mutations and subsequent cancer formation.

#### 1.2.1 Diet

The so-called western diet is characterised by higher proportions of processed foods and red meats, than other regional world diets, together with reduced dietary fibre and reduced protective phytochemicals found in fruits, vegetables, and whole grains compared to traditional diets[6]. With the latter having a potentially protective effect against colorectal cancer and the former being pro-oncogenic[7].

#### 1.2.2 Dietary fibre

The hypothesis that dietary fibre was important in the development of colorectal cancer was born out of global observations in the areas of low incidence (Africa, Asia) the consumption of fibre is greater than in the western world where the incidence is high[8]. The view that dietary fibre is protective against colorectal cancer, gained more credence in 2011 when the World Cancer Research Fund / American Institute for Cancer Research Report upgraded the association between fibre and risk of colorectal cancer from “probable” to “convincing”[9].

The presence of fibre in the diet leads to quicker colonic transit, dilution of colonic content and, most significantly, enhancement of bacterial fermentation, which leads to increased production of short-chain fatty acids (acetate, propionate, and butyrate), of which butyrate is recognised to have a

chemoprotective role by slowing growth and activating apoptosis in colon cancer cells[6,10]. Short chain fatty acids have been shown to interfere with numerous regulators of the cell cycle, proliferation, and apoptosis, such as β-catenin, p53, p21, Bax, and caspase 3 genes[11–13]. They have also been shown to have an anti-inflammatory effect decreasing the production of IL-6 and TNFα and also cycloxygenase 2 (COX-2) and inducible nitric oxide synthase (iNOS) gene expression[14,15]. Research conducted by the European Prospective Investigation into Cancer and Nutrition (EPIC) analysed fibre intake and incidence of colorectal cancer in more than 500,000 individuals. In doing so, they demonstrated a 40% reduction in the incidence of colorectal cancer when doubling fibre intake, with an even greater reduction in individuals who consume higher levels of fruit and vegetables[16].

A meta-analysis conducted in 2011, looking at 25 prospective cohort and nested case-control studies of dietary fibre or whole grain intake and incidence of colorectal cancer, concluded that a high intake of dietary fibre, in particular cereal fibre and whole grains, was associated with a reduced risk of colorectal cancer. The summary relative risk of developing colorectal cancer for 10 g daily of total dietary fibre (16 studies) was 0.90 (95% CI 0.86-0.94), for fruit fibre (n = 9) was 0.93 (95% CI 0.82-1.05), for vegetable fibre (n = 9) was 0.98 (95% CI 0.91-1.06), for legume fibre (n = 4) was 0.62 (95% CI 0.27-1.42), and for cereal fibre (n = 8) was 0.90 (95% CI0.83-0.97). The summary relative risk for an increment of three servings daily of whole grains (n = 6) was 0.83[17]. However, a Cochrane review conducted in 2002 did not demonstrate similar results and concluded that there was currently no evidence from RCTs to suggest that increased dietary fibre intake will reduce the incidence or recurrence of adenomatous polyps[18].

**1.2.3 B Vitamins- including folate ($B_9$)**

Fibre rich foods are frequently also high in B vitamins, including Vitamin $B_9$, folate. Results from the EPIC study were initially questioned, due to the confounding presence of folate, however subsequent analysis showed this to be independent[19]. B vitamins have a key role in DNA synthesis, repair, and methylation[20]. They have recently been shown to reduce the invasiveness of colorectal cancer by the activation of the Hedgehog Shh signaling pathway through promoter hypomethylation and stimulation of the nuclear factor κB (NF-κB) pathway[21]. This benefit has not been demonstrated when pharmacologically supplemented folate is given and is only present when taken via the diet[19]. In fact, excessive supplementation of folate in humans has been shown to increase colorectal cancer and adenoma formation[22].

Vitamin $B_6$, or pyridoxal phosphate, has also been shown to be important. A meta-analysis carried out in 2010 demonstrated the risk of colorectal cancer decreased by 49% for every 100-pmol/mL increase in blood Vitamin $B_6$ (RR 0.51; 95% CI 0.38-0.69)[23].

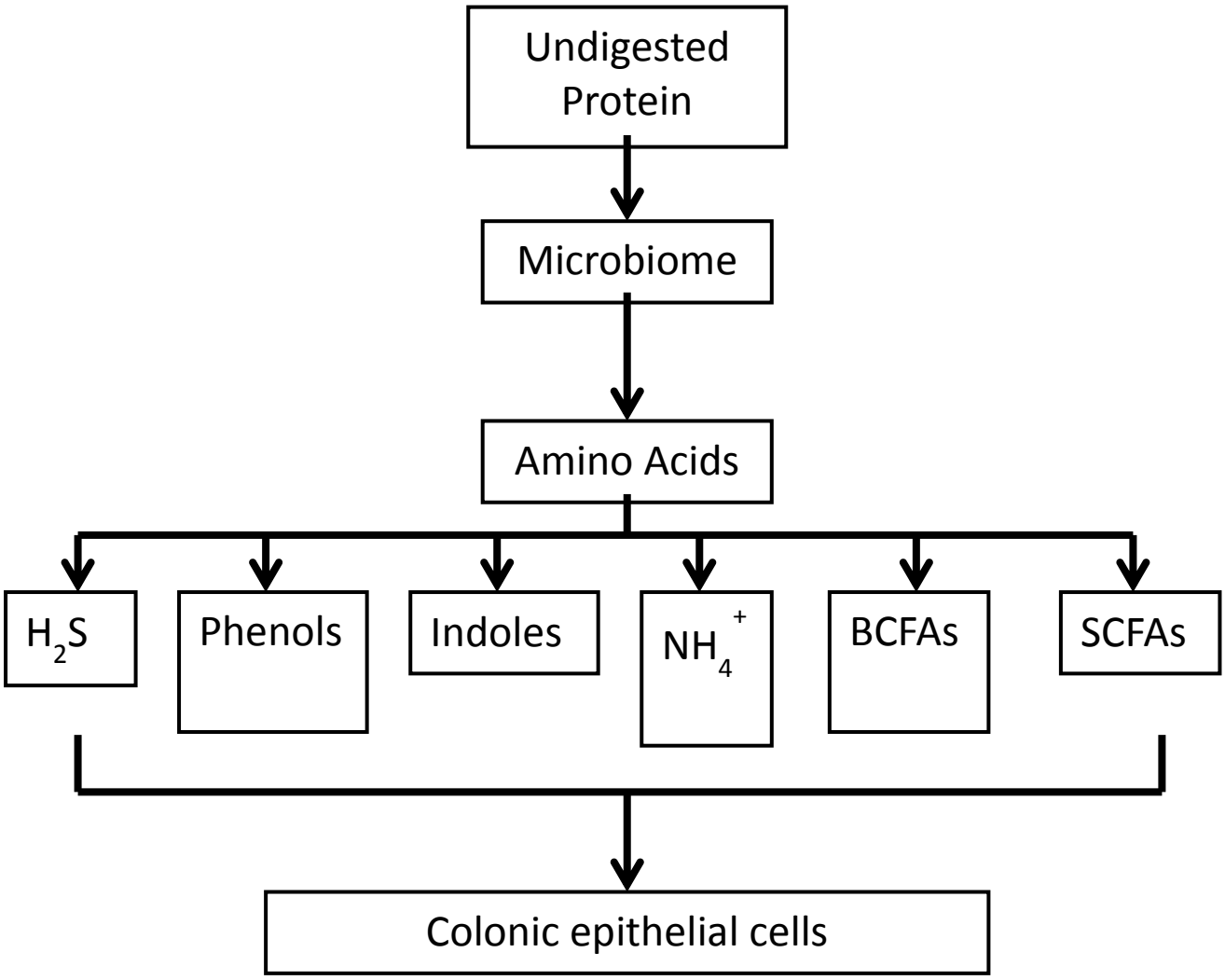
**1.2.4 Red and processed meat**

A meta-analysis from 2011 demonstrated further association between red and processed meat and colorectal cancer. This is particularly pertinent since the World Cancer Research Fund/American Institute of Cancer Research Report of 2007 upgraded the evidence for this to "convincing". The summary relative risk of colorectal cancer for the highest versus the lowest intake was 1.22 (95% CI 1.11-1.34) and the RR for every 100 g/day increase was 1.14 (95% CI 1.04-1.24). Non-linear dose-response meta-analyses revealed that colorectal cancer risk increases, approximately linearly, with increasing intake of red and processed meats up to approximately 140 g/day, where it begins to plateau[24]. This becomes particularly significant when considering the average daily intake in the USA is 128 g[25].

Potential reasons for the association between high red and processed meat intake and colorectal cancer risk include the content of the meat (e.g. protein, haeme) and compounds generated by the cooking process (e.g. N-nitroso compounds, heterocyclic amines). These factors can affect the large intestine mucosa with genotoxicity and metabolic disturbances, including expression of numerous cytokines (e.g. IL-6, IL-8, TNFα, NF-κB), leading to increased cytotoxicity and stimulation of an inflammatory response[26]. Furthermore the processing of red meat, including frying or grilling at high temperatures causes degradation of muscle creatinine and amino acids, resulting in the formation of numerous carcinogenic heterocyclic amines[6].

Increased bacterial fermentation (putrefaction) of undigested protein and production of bacterial metabolites derived from amino acids may affect colon epithelial homeostasis and renewal. This correlates with the fact that most colonic cancers are detected in the distal colon and rectum where protein fermentation occurs[27]. Animal models, with high protein diets, have demonstrated increased genetic damage in colon cells, as well as the level of faecal p-cresol (synonyms include 4-methylphenol, m-cresol and 4-cresol) , compared to that of rats fed a normal-protein diet[28,29].

**Figure 1: Schematic view of the metabolism of undigested proteins in the large intestine.**




Alimentary and endogenous undigested proteins undergo metabolism by the microbiota leading to the release of amino acids and further metabolism resulting in the production of a complex mixture of metabolic end-products that may enter into the colonocytes and exert deleterious effects when present in excess, adapted from Kim *et al*[27]. BCFAs= Branch chain fatty acids, SCFAs= short chain fatty acids.

However, this does remains an area of contention, with a series of conflicting studies. In a multi-ethnic cohort, with 165,717 participants, it was concluded that there was no role for meat in the aetiology of colorectal cancer[30]. A further analysis of more than 35 prospective studies failed to find a clear dose-response relationship between red meat intake and colorectal cancer[31].

**1.2.5 Smoking**

Unlike those factors arising from dietary exposures, there is little debate for the carcinogenic impact of smoking. The EPIC study demonstrated the clear link between tobacco smoking and colorectal cancer, in both current and previous smokers (RR 1.21), with most cancers being right sided[6,16].

The main carcinogens found in tobacco smoke are aromatic amines, nitrosamines, heterocyclic amines, and polycyclic aromatic hydrocarbons[6]. These substances, via their metabolism through cytochrome p450 enzymes, result in aberrant DNA and specific oncogenic mutation e.g. *KRAS*[32]. Nicotine also plays a key role through its activation of β-adrenoreceptors, which in turn trigger inflammatory and metastatic signalling through the COX-2, matrix metalloproteinase (MMP-2), and VEGF pathways[33].

**1.2.6 Alcohol**

Excessive alcohol has been linked to a number of gastrointestinal (GI) malignancies. Metabolism of alcohol via alcohol dehydrogenase, catalase and cytochrome p450, results in the production of acetaldehyde, which is a class 1 carcinogen and is responsible for chromosome damage[34]. Meta-analysis of 9 cohort studies, from 2014, showed not only a clear link between alcohol and colorectal cancer, but also a dose response relationship. Compared with non/occasional drinkers, the pooled RR was 1.03 (95% CI 0.93-1.15) for any, 0.97 (95% CI 0.86-1.10) for light (≤12.5 g/day of ethanol), 1.04 (95% CI 0.94-1.16) for moderate (12.6-49.9 g/day of ethanol), and 1.21 (95% CI 1.01-1.46) for heavy drinkers (≥50 g/day of ethanol)[35]. This dose response has also been demonstrated in meta-analysis considering intake below 50g/day[36]. This dose dependent relationship has been demonstrated on a cellular level, with an increase in DNA strand breaks in colonic mucosa cells with increasing ethanol levels[37].

A connection has also been described with adenomatous polyps, the precursor lesion to many colorectal cancers. Analysis of 30 studies, with 26 145 incident colonic adenomas, demonstrated an increase of 25 g per day of alcohol consumption led to an increased risk of colonic adenoma (summary relative risk=1.27, 95% CI 1.17-1.37). This association between alcohol consumption and adenoma formation applied equally to men and women, but did not apply to rectal adenomas[38].

In addition to the pro-oncogenic characteristics of alcohol, chronic exposure to alcohol results in deficiency of dietary B Vitamins, including folate. This provides a second potential mechanism for its association with colorectal cancer.

**1.2.7 Obesity**

Obesity, defined as a BMI >25kg/$m^2$, has been described as a contributing factor to many neoplasms, including colorectal cancer. A systematic review in 2013 showed this association, both when considering general obesity (BMI) and central obesity (waist circumference, WC). Approximately 9 million participants from several countries were included in this analysis. 41 studies on general obesity and 13 studies on central obesity were included in the meta-analysis. The pooled RRs of

colorectal cancer for the obese vs. normal category of BMI were 1.334 (95% CI 1.253-1.420), and the highest vs. lowest category of WC were 1.455 (95% CI 1.327-1.596)[39]. Another systematic review assessing colorectal cancer and adenoma indicated their higher prevalence in obese individuals, particularly in men. Compared with BMI below 23.0 kg/$m^2$, for BMI of 23.0–24.9, 25.0–27.4, and 27.5–29.9 kg/$m^2$ and BMI greater than 30 kg/$m^2$, the risk of colorectal cancer was 14, 19, 24, and 41 %, respectively[40].

Mechanisms that may account for the link between obesity and colorectal cancer risk include hyperinsulinemia, insulin resistance, inflammation, altered immune response, oxidative stress, as well as disturbances in insulin-like growth factors, adipokines, and sex steroids[41]. Adipose tissue is considered to be metabolically active and is responsible for the release of numerous cytokines, hormones and T-cell stimuli, all of which lead to chronic, low level inflammation, which leads to an increased risk of colorectal cancer[6]. Hormones specific to adipose tissue include leptin, adiponectin, and resistin are known to be carcinogenic. Moreover, they have been demonstrated to be raised in the serum of obese patients with colorectal cancer and adenomas when compared to controls[42,43].

**1.2.8 Exercise and physical activity**

The protective relationship of exercise and colorectal cancer has been demonstrated in a number of systematic reviews and meta-analyses some suggesting a 24% lower risk of colorectal cancer development in those undergoing regular exercise compared with those who have a sedentary lifestyle[44]. This benefit also extends to polyp development, with the same group showing 16% reduction in the incidence of colonic adenomas and 35% reduction in the incidence of large colonic polyps in the physically active group[45]. Other systematic reviews have suggested even greater impact. Cross-sectional studies have shown that regular physical activity (7 hours of brisk walking per week) lowers the risk of colon carcinoma by 40%. Physical activity also improves the outcome of patients already diagnosed with colorectal carcinoma: for example, patients with advanced disease have been found to survive significantly longer if they perform 4 hours of brisk walking per week, or the equivalent degree of physical exercise[46]. Excessive exercise has been shown to be harmful, as it induces high levels of oxidative stress and subsequent DNA damage; this harmful effect is not demonstrated with moderate exercise[47].

There has also been a description of the effects of exercise on the site of colonic cancers. Physical activity was related inversely to the risk of cancer at the proximal (RR 0.76, 95% CI 0.70-0.83) and distal colon (RR 0.77, 95% CI 0.71-0.83). Such a relationship could not be established for the rectum (RR 0.98, 95% CI 0.88-1.08)[48].

The protective nature of physical activity is likely to arise from increased insulin sensitivity, lower insulin levels, decreased body mass, and decreased adipose tissue volume, leading to a reduction of chronic inflammation. Physical activity causes the reduction of adipose tissue. Adipose tissue is known to produce TNFα, which in turn results in chronic low-grade inflammation and colorectal cancer. In people undertaking regular exercise, lower levels of TNFα and iNOS have been measured in the serum, with increased levels of IL-6, leading to enhanced immunity and also increased lipolysis in adipose tissue[49,50].

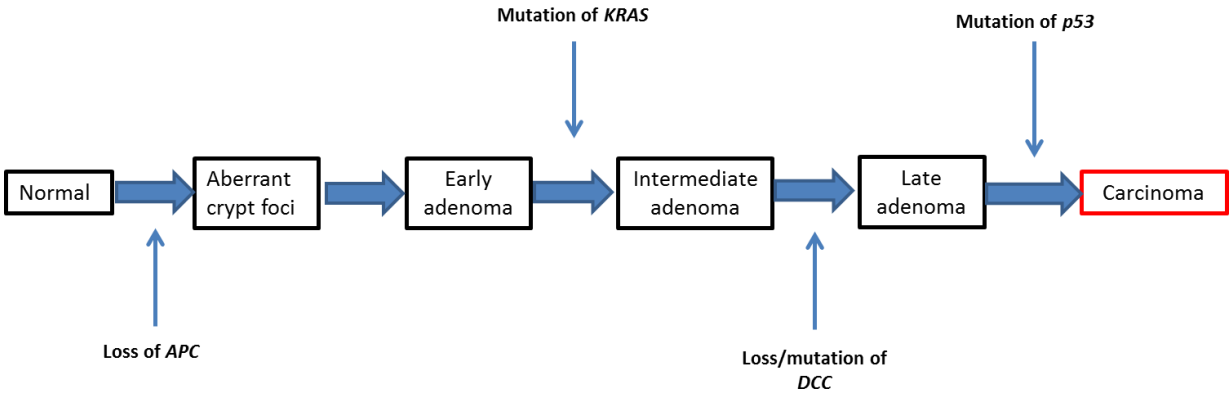
**1.2.9 Circadian rhythm**

The circadian timing system (CTS) controls several critical molecular pathways for cancer processes and treatment effects over the 24 hours, including drug metabolism, cell cycle, apoptosis, and DNA damage repair mechanisms[51]. There is clear variation in the mechanisms of cell proliferation, differentiation, apoptosis, and DNA repair during the day and night[52]. In fact, Hrushesky *et al*, showed a 50% increased risk of developing colorectal cancer in night-time shift workers[52]. Circadian clock genes have been described, these include *PER1*, *PER2*, *PER3*, and *CLOCK*[6]. These have been shown to up- and down-regulate oncogenes, such as *p53*, and to be key to cell signalling pathways, including c-myc, cyclin D, β-catenin, and VEGF[53,54].

### 1.3 Pathogenesis of colorectal cancer

Vogelstein first described a model of colorectal carcinogenesis arising from an accumulation of genetic alteration, now widely accepted[55]. The development of molecular genetic understanding has demonstrated these critical mutations underlying the pathogenesis of the sporadic and inherited forms of colorectal cancer (Figure 2). Recent genome wide studies have demonstrated up to 80 individual mutations per colorectal cancer, there is certainly a subset within this larger number of so called "drivers", of which there are fewer than 15[56,57]. These genes with a more prominent role include *APC, KRAS* and *p53*[58]. They have prominent roles in a number of inherited syndromes and also in the adenoma-carcinoma sequence.

**Figure 2: Adenoma-carcinoma sequence with associated genetic alterations leading to progression.**




Identifies the points at which specific genetic mutations and alteration occur, modified from Vogelstein[55].

For tumour progression to occur multiple genetic events, as outlined above, need to take place: genetic instability being key to this process. In colon cancer, at least 3 distinct pathways of genomic instability have been described, the chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathways[57].

#### 1.3.1 Chromosomal instability

Chromosomal instability (CIN) is observed in 65%–70% of sporadic colorectal cancers[57]. The consequence of this instability is loss of genetic heterozygosity (LOH) and aneuploidy. Chromosome segregation appears to be an important influence on CIN. If this is faulty, there is impaired segregation of chromosomes leading to an abnormal number of chromosomes being distributed to daughter cells[57]. Another proposed mechanism is abnormal centrosome number and function. This is pivotal during mitosis and chromosome separation, if there are too many centrosomes present then there is an unequal distribution of chromosomes to daughter cells and the potential for carcinogenesis[59]. The abnormalities that occur during CIN result in an impaired ability to repair

damaged DNA and a set of characteristic mutations in specific tumour suppressor genes and oncogenes which are crucial to the progression through the adenoma-carcinoma pathway.

**Table 1: Overall prevalence of genetic mutations in chromosomal instability–positive colorectal cancers.**

| Gene | Prevalence of mutation % | Function of gene product |
|---|---|---|
| KRAS | 30-50 | Cell proliferation, survival and transformation |
| CTNNB1 | 4-15 | Regulation of tumour growth and invasion |
| PIK3CA | 20 | Cell proliferation and survival |
| APC | 30-70 | Inhibition of Wnt signalling |
| TP53 | 40-50 | Cell cycle arrest, apoptosis |
| SMAD4 and 2 | 10-20 | Intracellular mediator of TGF-Beta |
| DCC | 6 | Cell surface receptor to netrin-1 |

Another key component of CIN is telomere dysfunction. Telomeres are DNA-protein complexes that have a protective role in preventing the fusion of chromosomal ends and disordered segregation. With each round of cell division a length of telomere is lost until they are unable to continue to function. Ordinarily this shortened telomere length would lead to programmed cell death. If this is incomplete and the cell manages to survive it will activate telomerase, in order to lengthen the remaining telomeres. In doing so it enters the breakage-fusion-bridge cycle that can continue for multiple cell generations and lead to dramatic genome reorganization[57]. Shorter telomere length has been demonstrated in both high grade dysplasia and established colorectal cancer[60]. Furthermore, an increase in telomerase activity has been shown to be associated with carcinoma. When comparing Dukes A to Dukes D carcinoma there is greater telomerase activity in Dukes D tumours[61].

### 1.3.2 Microsatellite instability (MSI)

The term microsatellite instability is applied, by international consensus, when >30% of the microsatellite marker panel is mutated[62]. Cancers with MSI account for approximately 15% of all colorectal cancers and 90% of colorectal cancers in patients with Lynch syndrome[63]. DNA mismatches, either insertions or deletions, can occur during replication, MMR protein complexes

bind to these defective areas and correct them, thus repairing the DNA sequence. If this MMR is defective then novel microsatellite fragments are created.

When MSI occurs in DNA mismatch repair genes (*MLH1, MSH2 MSH6* and *PMS2*) this gives rise to the Lynch Syndrome as described above. MSI tumours are typically located in the proximal colon, are poorly differentiated and have a mucinous appearance[62]. The frequency with which *MLH-1* is lost within colorectal cancer increases with age and by the age of 90 years, the gene is lost in 50% of tumours[64].

MSI can affect similar molecular pathways to those seen in CIN, but has interactions with different proteins. For example, unlike in CIN, MSI tumours have a normal *APC* expression, instead having a mutation in β-catenin meaning it is unable to interact with *APC*[65].

### 1.3.3 CpG island methylator phenotype (CIMP)

The CpG sites are regions of DNA where a cytosine nucleotide occurs next to a guaninenucleotide in the linear sequence of bases along its length. Throughout human evolution most of these sites have been lost, but about 1% of human DNA consists of short areas where CpG sites have escaped depletion[66]. Half of all human genes have their promoter regions embedded with these islands. Methylation of these areas, by DNA methyltransferase will silence the genes and such methylation will persist to daughter cells[67].

Cancer cells often have been found to have a global loss of methylation and a gain of methylation at the promoters of selected CpG islands, resulting in suppression of many genes, including tumour suppressor genes[68]. It has also been surmised that hypomethylation could lead to under-suppression on oncogenes. Observations relating to abnormal methylation, secondary to viruses, smoking and radiation have been described, but direct causality is unproven [67,69].

A subset of colorectal cancers are characterised by CIMP and they have a number of unique features. They are more common in the proximal colon, compared to the distal colon, and tend to occur in women[67]. Typically there are high levels of typical genetic abnormalities, namely of *p53* mutations and particularly of *KRAS*, such that nearly every CIMP-positive tumour has evidence of activation of the RAS oncogenic pathway[67,70].

### 1.3.4 The *APC* gene and colorectal cancer

The adenomatous polyposis coli (*APC*) gene is a key tumour suppressor gene in colorectal cancer. The *APC* gene product is a 312kDa protein that acts on various other proteins, including β-catenin. Through some of these interactions APC suppresses canonical Wnt signalling, which is essential for tumourigenesis. Further work demonstrates that APC plays roles in several other fundamental

cellular processes. These include cell adhesion and migration, organization of the actin and microtubule networks, spindle formation and chromosome segregation. Deregulation of these processes caused by mutations in *APC* is implicated in the initiation and expansion of colon cancer[71].

The Wnt signalling pathway is dependent upon β-catenin. APC protein is a major binding partner and regulator of the β-catenin protein[72]. In the absence of the APC gene, the regulation and destruction of β-catenin is lacking and so it accumulates. As a result, cytoplasmic β-catenin complexes with DNA-binding proteins of the TCF/LEF (T cell factor/lymphoid enhancer family) family, and translocates to the nucleus[71,72]. Dysregulation of a number of genes results, including up regulation of a number of oncogenes. Ultimately, this leads to the development of colorectal cancer. This pathway has been shown to be fundamental in inherited and, more so in, sporadic colorectal cancer. It is thought that the *APC* mutation is the initiating step in the adenoma-carcinoma pathway, and perhaps the initiating step in up to 80% of all sporadic carcinomas[58].

**1.3.5 The *KRAS* gene and colorectal cancer**

*KRAS* is mutated in 30%–50% of colorectal cancers[73]. Under normal conditions growth factors, binding to their cell surface receptors, activate guanine exchange factors (GEF), such as SOS (son of sevenless) that are attached by the adaptor protein GRB2 (growth-factor-receptor bound protein 2). SOS stimulates the release of bound guanosine diphosphate (GDP) from RAS, and it is exchanged for guanosine triphosphate (GTP), leading to the active RAS-GTP conformation[57]. This in turn activates downstream signalling. As a result of the mutation in the *KRAS* gene, the usual conversion back to RAS and GDP does not occur, meaning there is continual downstream activation by RAS-GTP, with sustained proliferation signalling in the cell.

*KRAS* mutations are present in hyperplastic lesions with little or no malignant potential[74]. However they have also been shown to increase in frequency in enlarging adenomatous polyps: up to 50% of adenomas >1cm showing *KRAS* mutations, whilst <10% of adenomas smaller than 1cm have them[55].

The presence of *KRAS* mutation and its association with prognosis has been described with conflicting comments. Confusion arises from the conflicting genetic factors associated with *KRAS* mutations. Principally, *KRAS*-mutated colorectal cancer is less likely to exhibit MSI and is almost never *BRAF*-mutated[75]. Tumours with high levels of MSI have been shown to have a better prognosis, whilst *BRAF*-mutated cancers have a poorer prognosis[76,77]. A study did show a poorer prognosis for those with *KRAS* mutation, but only in those with Dukes C and only in those with the

*KRAS* p.G12V mutation[78]. The presence of *KRAS* has also been shown to be a negative predictor in response to therapy with anti-epidermal growth factor receptor antibodies, such as cetuximab[79].

### 1.3.6 The *p53* gene and colorectal cancer

The *p53* tumour suppressor gene encodes a nuclear phosphoprotein with the ability to bind directly to DNA and act as a transcriptional activator[80]. In most tumours, the two *TP53* alleles are inactivated, usually by a combination of a missense mutation that inactivates the transcriptional activity of p53 and a 17p chromosomal deletion that eliminates the second *TP53* allele[81]. The role of the *p53* gene is to arrest the cell cycle in response to the damage in order the facilitate repair or the initiation of apoptosis[82]. In colorectal cancer, *p53* mutations play a pivotal role in the transformation of large adenomas into invasive carcinoma, as they promote continued growth, survival against various stresses and invasive properties[58,83].

### 1.3.7 Growth factor pathways

The activation of growth factor pathways is common in colorectal cancer, some of these are described in this section.

#### 1.3.7.1 Aberrant regulation of prostaglandin signalling

Adenoma development relies on the critical step of activation of prostaglandin signalling[81]. Up-regulation of COX-2 leads to overproduction of prostaglandin $E_2$, this is strongly associated with colorectal cancer[84]. Increased levels of prostaglandin $E_2$ can also result from the loss of 15-prostaglandin dehydrogenase (15-PGDH), the rate-limiting enzyme in catalysing degradation of prostaglandin[81]. This specific change is seen in up to 80% of adenomas and cancers[84]. There have been a number of clinical trials that have demonstrated a reduction in development of new adenomas and regression of existing adenomas in the presence of COX-2 inhibitors[81,85,86].

#### 1.3.7.2 Epidermal growth factor receptor

Epidermal growth factor (EGF) is a soluble protein that has trophic effects on intestinal cells. Clinical studies have supported an important role of signalling through the EGF receptor (EGFR) in a subgroup of colorectal cancers[81]. As previously described, advanced colorectal cancers with tumour-promoting mutations of these pathways, including activating mutations in *KRAS* and *BRAF,* do not respond to anti-EGFR therapy[79,81].

#### 1.3.7.3 Vascular endothelial growth factor

Vascular Endothelial Growth Factor (VEGF) is produced both as a result of injury and normal growth in order to promote angiogenesis. These angiogenic pathways have been shown to be key to the

growth and maintenance of colonic tumours and provides new potential targets for therapy, for example bevacizumab in metastatic colorectal cancer.

### 1.4 Health screening and colorectal cancer

Often described as Wilson's criteria, the principles of screening were defined in 1968 by the World Health Organisation[87]. These criteria are still applicable today and form the basis of modern health screening. They consist of 10 key points:

1. The condition should be an important health problem.
2. There should be a treatment for the condition.
3. Facilities for diagnosis and treatment should be available.
4. There should be a latent stage of the disease.
5. There should be a test or examination for the condition.
6. The test should be acceptable to the population.
7. The natural history of the disease should be adequately understood.
8. There should be an agreed policy on whom to treat.
9. The total cost of finding a case should be economically balanced in relation to medical expenditure as a whole.
10. Case-finding should be a continuous process, not just a "once and for all" project.

Clinical screening strategies are essentially of two types: those that involve direct investigation of the patient, often with an invasive procedure such as colonoscopy or smear testing, and those based on laboratory analysis of patient specimens, such as urine, faeces or peripheral blood[88]. As above, Wilson's criteria have been used to define health screening, however to be truly utilised, a screening tool must satisfy 4 key points;

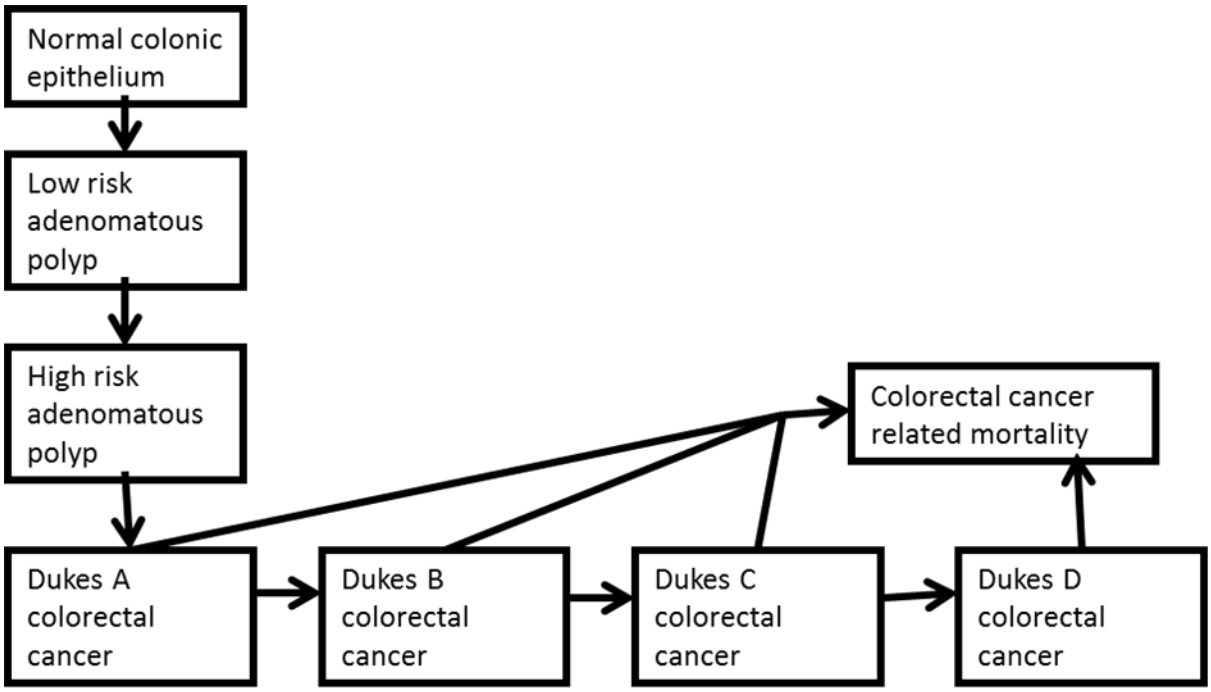
1) it must have acceptable performance in terms of positive and negative predictive value for the target population;
2) a reduction in mortality from the disease;
3) acceptable levels of patient compliance;

4) and low enough cost to be affordable within the context of the specific health-care system[88].

It is unrealistic to expect a completely ideal screening tool, but around the world a combination of non-invasive faecal occult blood testing (FOBt) followed by colonoscopy in those with an initial positive FOBt is an example for colorectal cancer.

The progression of normal mucosa to adenomatous polyp and then to colorectal cancer make colonic cancer a suitable screening target.

**Figure 3: Simplified colorectal cancer natural history progression model**




Adapted from Tappenden *et al*[89].

It has been stated that over 95% of colorectal cancer cases would benefit from surgical treatment, a potential cure, if diagnosed early or at the stage of the premalignant polyp[88]. It is estimated that in developed countries 40-50% of the population will have at least one colonic polyp.

A report by the European Commission in 2008 revealed that only 12 of its then, 22 member states, had organized population-based colorectal cancer screening programs in place or due to be activated. Following this report member states were urged to ensure fair access to preventive services and cancer screening programs for all citizens[90]. In 2010, new guidelines were published by the EU relating to colorectal cancer screening and diagnosis quality levels[91]. Not only did this include guidance on implementing and maintaining a screening programme but also included

encouragement of member states to seek improved primary prevention, such has improving dietary and environmental factors that lead to the development of colorectal cancer. The colorectal cancer screening method envisaged by European Council directives is based on a faecal occult blood test (FOBt) offered to 50–74 year olds, but different strategies are adopted in different countries. The UK completed its roll out in 2010.

European-wide guidance for bowel cancer screening suggests a faecal occult blood test (FOBt) between the ages of 50-74, followed by a colonoscopy if FOBt is positive, as the gold standard[90,92]. Two forms of FOBt (immunological e.g. OC-Sensor and guaiac e.g. Hemoccult-II) are currently available and used for screening purposes worldwide[93].

### 1.4.1 Faecal occult blood testing

A guaiac-based Faecal Occult Blood Test (gFOBt) uses the chemical alpha-guaiaconic acid (reagent derived from wood resin of Guajacum trees) to detect haem in stool. After stool is smeared on to the device hydrogen peroxide is added. As haem contains pseudoperoxidase there is a reaction which converts guaiac to the colour blue, giving an indication there may be blood in the stool. Foods which can give a false positive result include red meat, carrots, potatoes and figs. The guaiac-based faecal occult blood test usually picks up a daily blood loss of approximately 10 ml. The gFOBt currently used in the England Bowel Cancer Screening Programme (BCSP) has a sensitivity of 36% and a specificity of 94% for the detection of colorectal cancer when colonoscopy is used as the gold standard for cancer detection[94–96].

Immunological FOBt is an instant and objective test which utilizes immuno-chromatography to detect occult blood loss in stool. It is specific to human haemoglobin and only reacts with the intact haemoglobin molecule. It therefore requires no dietary restriction and it only identifies pathology limited to the colon. With all this in mind it has a lower false positive rate, thus reducing unnecessary colonoscopy[97]. The detection “cut-off” can be adjusted with iFOBt thus allowing it to be more adaptive to the needs of an individual population of programme[98]. In addition, testing may be performed on a single stool sample and integrated into the clinical examination providing an instant result which can be used as an adjunct to clinical assessment[99].

Faecal DNA testing is a promising development for point of care testing and thus screening, in conjunction with the gold standard colonoscopy. This essentially revolves around the analysis of DNA alterations present in tumour cells exfoliated into faeces[100]. The normal colonic epithelium renews itself daily, thus resulting in large numbers of exfoliated cells being present in faeces. It has been shown that epithelial exfoliation is more pronounced in the presence of colorectal cancer, with

colon cancers accounting for <1% of the total epithelial surface area but between 14-25% of the DNA retrieved from faecal analysis[101]. This may be because cancer cells are able to survive much longer due to the inhibition of apoptotic and cell death processes

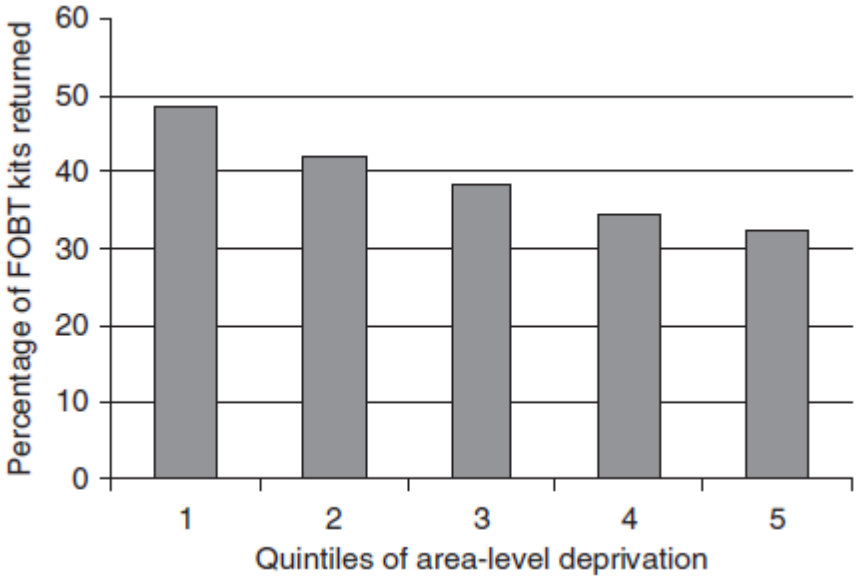
### 1.4.2 Adherence and compliance with colorectal cancer screening

Greater levels of population participation in colorectal cancer screening programs are associated with reduced mortality and greater cost-effectiveness[102]. The European Commission has previously set 45% as an acceptable uptake level[91], but greater than this is certainly desirable; similar American guidance sets the level at 75%. Across Europe, within countries and within areas covered by the same UK BCSP, uptake varies greatly. When assessed in boroughs of London, (UK), gFOBt return rate varied from 32 and 49%[103].

Uptake rates between men and women have been extensively studied, with contrasting results. Male participation appears to be greater, but this does vary between countries. Others have suggested that female participation, particularly in Europe and Australia is greatest[90]. In the UK female uptake is 4.8% greater than males, with the greatest difference being seen in London[5]. However, overall there is probably little significant difference between the two sexes[102].

Socio-economic status (low income, unemployment, low education and area of residence) appears to have the most significant impact on uptake to colorectal cancer screening programmes [90,103]. Von Wagner *et al* showed a direct correlation between socio-economic groupings and uptake to the colorectal cancer screening in London, with those living in areas of social deprivation having lower participation rates[5]. This is a persistent finding of many studies [90].

**Figure 4: FOBt return by quintile of area-based socio-economic deprivation**




Taken from von Wanger *et al* (n=401 197)[103]. (Reproduced with permission from British Journal of Cancer).

Married people are more likely to participate than those who are single, whilst smokers are less likely to participate than non-smokers[90].

Given the low participation in colorectal cancer screening programs, despite the clear medical benefit, it is important to understand the barriers to screening to develop successful alternative approaches. A series of studies have described such barriers, these include factors specific to the tests themselves, such as embarrassment, reluctance to handle stool, fear of the procedure, or inconvenience, as well as broader factors such as lack of access to care, limited knowledge of screening and a lack of physician recommendation[104]. A study from Germany showed that patient choice can be key to improving compliance: they offered a stool or blood test prior to colonoscopy and found well in excess of 80% preferred blood testing[105].

Patient knowledge is a key feature when considering compliance with bowel cancer screening. This includes knowledge of the procedure and intended benefits. Studies have posed questions including those about risk factors for developing colorectal cancer, incidence, prognosis, age-related risk, warning signs or symptoms, and knowledge about recommended colorectal cancer screening tests. There is huge international variation in patient knowledge, which may account for the variation in uptake. In the USA some studies have reported knowledge of colorectal cancer and screening was in excess of 80%, whilst in the UK it has been reported to be as low as 30%[104,106].

The Health Belief Model (HBM) can be used to understand patient participation in screening programmes. This model theorizes on people's beliefs regarding the risk for a disease or health problem, and according to their perceptions on the benefits of taking actions to avoid it and analyzes their readiness to take action. Considering this, people with negative attitudes such as embarrassment, anxiety, disinterest, fear of cancer or screening, lack of time, feeling healthy, apprehensions about the bowel preparation, handling stool for FOBt, laxatives or insertion of a tube, and discomfort are more reluctant to participate in screening programmes[107–109]. A Spanish study, from 2009, showed that embarrassment was the leading barrier to participation in the bowel cancer screening programme, this included stigma of handling stool[110].

**1.4.3 Impact of bowel cancer screening**

Randomised trials have shown that screening for bowel cancer using guaiac-based faecal occult blood tests (gFOBts) can reduce mortality by 16% in people offered screening and 25% in those accepting it[111,112]. Other studies have shown that gFOBt screening followed by colonoscopy in gFOBt-positive patients produced a 33% reduction in colorectal cancer mortality with gFOBt performed yearly and a 21% reduction with gFOBt performed at 2-year intervals; incidence rates fell by 20% and 17%, respectively[113,114]. Zorzi *et al* demonstrated a superior reduction in mortality when faecal immunochemical test FIT based screening is adopted compared to gFOBt[115]. Analysis of the economics has shown it to be beneficial, with a cost per quality-adjusted life-year gained of < £3000 for gFOBt screening[89].

Earlier lesion detection is a recognised effect of colorectal cancer screening, both in terms of established cancers and polyps. A Cochrane review by Hewitson *et al* showed a significant reduction in Dukes D and increased Dukes A detection[112]. Other such studies report an increased diagnosis of Dukes A lesions with colonoscopy and a corresponding reduction of Dukes C lesions in subsequent screening rounds[116,117]. Polyps are also detected at an earlier stage, when they can be considered lower risk. This becomes a significant feature when considering the subsequent risk of colorectal cancer and associated death after removal of a high risk polyp. Loberg *et al* assessed 40,826 people following polypectomy over a median follow-up of 7.7 years (maximum, 19.0): a total of 398 deaths from colorectal cancer were expected and 383 were observed, for an SMR of 0.96 (95% CI 0.87-1.06) among patients who had had adenomas removed. Colorectal-cancer mortality was increased among patients with high-risk adenomas (expected deaths, 209; observed deaths, 242; SMR, 1.16; 95% CI 1.02 to 1.31), but it was reduced among patients with low-risk adenomas (expected deaths, 189; observed deaths, 141; 95% CI 0.63 to 0.88)[118].

On the whole the literature suggests an overall reduction in mortality from colorectal cancer as a result of bowel cancer screening programmes, however some studies have not shown a reduction in all-cause mortality and there is active debate as to whether disease-specific or direct cancer death is the most appropriate endpoint[112].

**1.4.4 Bowel cancer screening programme in the UK**

In the UK screening takes place on a biennial basis using non-rehydrated guaiac-based faecal occult blood tests, followed by colonoscopy with participants being aged 60-69 years. Whilst in place, this programme has been shown to reduce mortality from colorectal cancer, increase adenoma detection rates and be cost effective[119,120]. For the UK, the proportion of subjects returning kits with positive test results was 2.0% and was higher in men (2.5%)[5]. From its initiation until 2008, 17518 positive FOBts have been resulted in investigation, 83% of the total positive tests. Of this 17518, 98.1% had a colonoscopy. 1772 subjects were recorded as having a colorectal cancer (10.1%), together with 4,777 having either intermediate or high risk polyps requiring further intervention. A total of 71.3% of the cancers were polyp cancers or Dukes A or B and so potentially curable. 28.7% of the cancers were found in the rectum or recto-sigmoid colon and, overall, 77.3% were recorded as being left-sided colorectal cancers while only 14.3% were recorded as being in the right colon. As expected right-sided cancer was more commonly found in women (19.2%) than in men (12.2%) and conversely rectal cancer was less common in women (20.3%) than men (28.5%)[5].

A one-off flexible sigmoidoscopy, known as bowel scope screening, is a new part of the UK BCSP. It began its nationwide roll out in 2013. It involves a once-only flexible sigmoidoscopy in men and women once they turn 55 years of age. Atkin *et al* demonstrated that such a programme could reduce the incidence of colorectal cancer by 33% and mortality by 43%, over an 11-year follow-up. Moreover the incidence of distal cancers was reduced by 50%[121].

These early results indicate that the BCSP in England is on track to match the 16% reduction in colorectal cancer mortality found in the randomised trials of gFOBt screening[5,112]. Replacing the gFOBt with another method, such as FIT, DNA analysis or novel technique currently in development may improve this further.

By identifying colorectal cancer at an earlier stage and advanced pre-malignant neoplasia, it is clear that bowel cancer screening improves long term outcomes from colorectal cancer. Limiting colonoscopy in those with a normal colon, thus a false positive FOBt, is a key issue for future service provision.

**1.5 Hepatocellular carcinoma**

Primary liver cancer (hepatocellular cancer, HCC) is the third largest contributor to cancer mortality in the world, with more than half a million people worldwide diagnosed each year[122]. Liver cancer is the fifth most common cancer in men and the seventh in women. Most of the burden of disease (85%) is borne in developing countries, with the highest incidence in regions where infection with hepatitis B virus (HBV) is endemic[123]. Over the last 3 decades the incidence of HCC has been steadily increasing, this has been mirrored by an increasing mortality rate. Age-standardized rates are particularly high in eastern and south-eastern Asia (over 20/100,000 men and over 10/100,000 women) and middle and western Africa (15-20/100,000 men and about 8-19/100,000 women); in most high-income countries, including the Americas, Australia, western and northern Europe, rates are below 7.5/100,000 men and below 2.5/100,000 women, while intermediate rates (around 10/100,000 men and 3/100,000 women) are observed in southern Europe[124].

Mortality and 5-year survival from HCC is variable worldwide, but generally would be considered to be poor. This may stem from the variation of underlying aetiology, surveillance programmes, economics factors and mortality reporting, along with a long list of other potential factors. In Europe, 5-year survival is reported to be approximately 12%, but there is significant variation between its member states[125]. The highest overall mortality rates from primary liver cancer in men were in France (6.2/100,000), Spain (4.9), Austria (4.3), and Italy (4.0), while the lowest ones were in Sweden (1.1), the Netherlands (1.2), the UK (1.8), and Denmark (1.9)[124]. However, in recent years, there has been a rising trend in mortality in the UK and Germany, whilst rates are beginning to fall in France and Italy. Consequently, mortality rates are becoming more uniform across Europe. In lower income countries, around the world, the 5-year survival has been reported to be as low as 5%, interpretation needs to be with some caution due to the questionale reliability of diagnosis and death certification.

HCC typically occurs on a background of liver cirrhosis (80-90%) and most commonly in the presence of viral hepatitis, namely hepatitis B and C (HCV). After viral aetiologies, alcoholic liver disease and non-alcoholic fatty liver disease are the most common underlying causes, with less common causes such as hereditary haemochromatosis, $alpha_1$-antitrypsin deficiency, autoimmune hepatitis, some porphyrias, and Wilson's disease being described[123].

Worldwide, chronic HBV infection accounts for approximately 50% of all cases of hepatocellular carcinoma and virtually all childhood cases[123]. Infection with HBV can lead to the development of HCC in the absence of cirrhosis, but this only occurs in approximately 20% of patients with HBV-related HCC. The risk of developing HCC in the presence of HBV is further increased if there is co-

infection with HCV, alcohol consumption, increasing age and high circulating HBV DNA levels[123,126].

HCCs develop as small nodules. The majority of their growth takes place in an asymptomatic phase which may be years in length, the median time of size doubling being six months. The major factors contributing to survival are the degree of the underlying liver disease and the size of the tumour at diagnosis[127]. According to specific clinical parameters the Childs Pugh score can be calculated in order to define the severity of cirrhosis. Childs Pugh A cirrhosis has an 85% 2-year survival rate, compared to an 80% 2-year mortality for Childs Pugh C. Similarly, as size increases so does mortality if cases are left untreated. This all suggests that early diagnosis and treatment will have an impact upon mortality rates, thus implying the need for screening of at-risk groups as a potential measure. Moreover, cirrhosis is a progressive disease that affects patient survival. The presence of cirrhosis then influences the chances of receiving an effective treatment, thus making an early diagnosis of HCC even more crucial.

Decision analysis and cost–effectiveness models suggest that an intervention is considered cost-effective if it provides gains of life expectancy of at least 3 months with a cost lower than approximately US$ 50,000 per year of life saved[128,129]. This cost effectiveness can be further improved by identifying those at low risk and those whose liver disease is too advanced. Such a strategy has led to the recommendations by the European Association for the Study of Liver (EASL), stating that the following patients should be offered HCC surveillance:

-Cirrhotic patients, Childs-Pugh A and B

-Cirrhotic patients, Childs-Pugh C awaiting transplant

-Non-cirrhotic HBV carrier with active hepatitis or family history of HCC

-Non-cirrhotic patients with chronic hepatitis C and advanced liver fibrosis

**1.5.1 Screening and HCC**

In order to effectively define the role of screening in HCC, an ideal randomized, controlled study would compare a group of at-risk subjects undergoing surveillance to a group that does not undergo surveillance. In order to increase efficiency, only the patients with the highest risk of HCC should be included. Those in the arm undergoing surveillance should receive ultrasound every six months[130]. Disease stage at diagnosis, treatment efficacy and mortality outcomes could then be accurately compared, in order to ascertain whether surveillance reduces the mortality from HCC. Another strategy to assess the impact of surveillance is to study cohorts of screened patients who develop

HCC compared to unscreened patients who develop HCC, however this process is open to lead-time bias making the results difficult to interpret and generalise[130]

The results from one of the first randomised HCC surveillance studies, conducted in China, were published in 2003. It compared surveillance with alpha-fetoprotein (AFP) to no surveillance. In the study, an AFP cut-off of >200g/ml was used. Subjects with such a result were then subjected to ultrasound assessment. As one may expect, there were a greater number of tumours diagnosed in the surveillance arm, together with more tumours being diagnosed at an earlier stage. Despite this the study failed to demonstrate a survival benefit[131]. This study was widely criticised as the treatment given following a diagnosis of HCC was inconsistent and in many cases not completed, thus making comparison of outcomes difficult. Other such earlier studies, also conducted in China, showed a reduction in mortality with combined AFP and ultrasound screening, in one instance up to a 37% reduction[132].

It is very unlikely that such an ideal study, to compare surveillance to no surveillance in high risk groups, will be conducted. Not only would it be unethical but research into this matter showed categorically that patients will not consent to such a study [133].

In the absence of this "ideal" study there is some evidence to suggest a benefit of screening for HCC[134]. In 2014, Yeh *et al*, from Taiwan, undertook a study in a population selected by establishing a risk score. The risk score was constructed from a panel of blood tests, and a history of diabetes or chronic viral hepatitis. Those with a risk score above a cut-off were invited to participate in a one-off screening protocol. The mortality in this group was then compared to the mortality of a similar group who did not respond to the invitation to be screened, and to the general population not invited to participate in the programme. This study did demonstrate a reduction in mortality from HCC for the surveillance group[134]. The high prevalence of viral hepatitis in this area could make it difficult to generalise these results to the UK population. A number of studies have shown that screening results in a greater detection of earlier stage tumours. These are potentially curative by resection or transplant, but have failed to show a significant difference in survival[135,136].

As described above, cost effectiveness is another way to analyse the impact of screening for HCC. There are a number of studies that look at this aspect but there is little homogeneity between them, with different surveillance techniques and intervals, different intervention after diagnosis and different costs. The diversity in the underlying aetiology of cirrhosis and subsequent disease course greatly affect any cost effectiveness analysis. What they do have in common is that they show that the cost effectiveness of surveillance is critically dependent on the incidence of HCC, but virtually all

show surveillance is cost effective[130]. On balance, surveillance appears to be cost effective if the incidence of HCC exceeds 1.5-2%/year[137,138] which is consistent with AASLD recommendations.

Andersson *et al*, in 2008, demonstrated the cost effectiveness of ultrasound. They also showed that the use of CT and MRI for screening is not cost effective, with an incremental cost effectiveness ratio of more than $100,000 to $300,000, respectively[139].

**1.5.2 Evidence for surveillance modalities and intervals in HCC**

Until relatively recently screening has involved the use of AFP, either alone or in combination with another modality such as ultrasound. Much of the data suggest that AFP contributes little to the detection of HCC when it is used in combination, with a sensitivity for small HCCs of approximately 60-65%[140,141]. It also carries a significant false-negative rate, with approximately 40-50% of people with a diagnosis of small HCC having a normal AFP[130,142]. AFP has an unacceptable false-positive rate as it is elevated in patients with active hepatitis and cirrhosis without HCC. There is also clear evidence that shows a higher AFP is more likely to be associated with advanced size, thus limiting its role as a screening tool for early cancer [143].

The application and the sensitivity of ultrasound for the detection of HCC is highly operator dependent: it may be 80% in expert hands, however, when a more generalised approach is assessed, the sensitivity for early HCCs is 65% i.e. similar to AFP and other biomarkers. A recent meta-analysis including 19 studies showed that US surveillance detected the majority of HCC tumours before they presented clinically, with a pooled sensitivity of 94%. However when further analysis looked at early lesion detection the sensitivity fell to 63%[144]. Operator dependence was clearly highlighted by a Japanese study of 1432 patients: careful US surveillance performed by highly skilled operators resulted in an average size of the detected tumours of 1.6 ± 0.6 cm, with less than 2% of the cases exceeding 3 cm[145]. Ultrasound is a very appealing screening modality because of the absence of risks, non-invasiveness, good acceptance by patients and relatively moderate cost. There was hope that contrast assisted ultrasound would improve the detection of early HCC by reducing the inter-operative dependency, but this was shown not to be the case[146].

The performance of CT and MRI as surveillance tools for HCC has not been evaluated and so no data exist [129,130]. There are significant issues related to using CT and MRI for HCC screening. Cost and accessibility are two such issues, there needs to be a much greater infrastructure in place to use these modalities, which significantly increases cost and reduces accessibility. There is also a significant false positive rate with CT and MRI scanning for surveillance, due to intra-hepatic shunts, dysplastic nodules and cirrhotic nodules. EASL guidelines suggest that these techniques should be

also considered when obesity, intestinal gas, and chest wall deformity prevent an adequate US assessment. Even in these circumstances, radiation risk due to repeated exposure to CT scan and high cost of MRI make debatable their use in long-term surveillance[129].

### 1.5.3 Screening schedule

The ideal interval of surveillance for HCC should be dictated by two main features: rate of tumour growth up to the limit of its detectability, and tumour incidence in the target population. The doubling time of HCC suggests the screening interval time should be 6 months. Some studies have shown variation of this, in relation to the differing underlying aetiologies, others show 6 months to be appropriate. Han *et al* compared 6-monthly to 12-monthly and demonstrated that survival was better with 6-monthly surveillance than with 12-monthly surveillance[147,148]. Six-monthly surveillance has been compared to 3-monthly: it did not reduce mortality, but did lead to a higher number of false-positives and a greater number of investigations to evaluate these[149]. Meta-analysis of prospective studies has shown that the pooled sensitivity of US-based surveillance decreases from 70% with the 6-month program to 50% with the annual program[144]. Cost-effectiveness studies have shown that 6-monthly ultrasound based surveillance improves quality-adjusted life expectancy at a reasonable cost, further supporting the use of ultrasound as the preferred imaging modality in HCC surveillance[139].

### 1.5.4 Current recommendations

There is still controversy around the benefit and recommendation of HCC surveillance. A recent systematic review concluded that there was insufficient evidence to recommend HCC surveillance[150]. However the majority of the literature concludes that, despite relatively poor evidence, it should take place[129,130,151]. All surveillance programmes bring with them the possibility of harm to the patient. This may be psychological, related to anxiety waiting for results every time surveillance is done, or it may be harm resulting from investigation of false-positives. In the case of HCC the potential harms are those mentioned above, plus the possibility of treating a cancer that might never have caused any clinical symptoms, plus the harm related to complications of treatment. Indolent HCCs are very rare and in fact malignant nodules detected on ultrasound that are not treated grow and become obviously malignant within the space of a few years[130].

Current EASL guidelines are:

- Patients at high risk for developing HCC should be entered into surveillance programmes (High risk groups seen in figure 6)

- Surveillance should be performed by experienced personnel in all at-risk populations using abdominal ultrasound every 6 months. A shorter period of 3-4 months should be employed where a nodule of less than 1 cm has been detected or in surveillance after local resection.
- Patients on the waiting list for liver transplantation should be screened for HCC in order to detect and manage tumour progression and to help define priority policies for transplantation.

Such an approach is generally supported by the majority of the available literature and is routine practice in the UK. Screening will, ideally, identify HCC at an earlier stage, when curative intervention can be offered. This has been defined by the Barcelona Clinic Liver Cancer staging (BCLC)[129,152]. This staging system scores disease as 0, A, B, C or D according to lesion size, number of lesions and underlying composite liver disease. It then defines 0 and A as potentially curative, whilst B-D are non-curative and should receive treatments such as radiofrequency ablation (RFA) or transarterial chemoembolisation (TACE). Therefore an ideal screening tool would identify those classified as 0 or A in order to offer potentially curative treatment.

**1.6 Pathogenesis of HCC according to aetiology**

The variety of the underlying aetiologies of HCC means that the disease is heterogenous, both from a molecular and a clinical standpoint. This is mirrored on a global scale with HBV being prominent in eastern countries, whilst alcohol, non-alcoholic steatohepatitis and HCV are dominant in the west[123]. In >90% of cases HCC will develop on a background of cirrhosis, therefore the severity of the underlying cirrhosis will influence molecular and clinical factors. These changes drastically alter the matrix and microenvironment of the liver[153,154]. The risk of HCC in patients with liver cirrhosis depends on the activity, duration and the aetiology of the underlying liver disease. Coexistence of aetiologies, e.g., HBV and HCV infection, HBV infection and aflatoxin B1, HCV infection and alcohol, or HCV infection and liver steatosis, increases the relative risk of HCC development[155]. Moreover, the aetiology of the non-cirrhotic HCC will be different as this altered matrix and microenvironment is not present.

The main causative agents—HBV, HCV and aflatoxin B1 (AFB)—which together are responsible for about 80% of all HCCs in humans, leave 'molecular marks' on hepatocytes that enable the causes of individual HCCs to be determined accurately in many instances[156]. HCC, like other cancers, is a disease of the genome and is defined by malignant hepatocytes accumulating somatic genetic alterations that combine mutations in both pro-oncogenes and tumour suppressor genes[157]. This

process can take place over a long period, up to 30 years, after chronic HBV or HCV have been diagnosed. Pre-neoplastic cells may take between 10-30 years to develop, this is followed by dysplasia and ultimately carcinoma[156].

### 1.6.1 Pathogenesis of HCC related to HBV

An estimated 300–400 million individuals worldwide are chronically infected with HBV. Prevalence rates range from 0.1 to 1% of the general population in North America and western Europe, to up to 20% in southeast Asia and parts of Africa[155]. Studies have shown that areas with a high prevalence of HBV carriage also have a high incidence of HCC[158]. It is the leading aetiology in Asian and African countries. Compared to those not infected with HBV the life time risk of HCC is 100-fold greater in those infected, with approximately 40% of males infected with HBV developing HCC[155]. Unlike many of the other causes of HCC, HBV infection can result in HCC development in the absence of cirrhosis[159], but in the majority of cases there is the underlying chronic inflammation, fibrosis and cirrhosis .

One specific pathogenic feature of HBV infection is the integration of viral DNA into human DNA. Integration of HBV DNA is not part of the viral life cycle, but rather occurs as an epiphenomenon of HBV replication[155]. Overall, at least 70-80% of HCC due to HBV infection show clonal integration of the viral genome[160,161]. It would appear that the site of HBV DNA insertion is somewhat random, potentially explaining why the majority of people with chronic HBV infection do not develop HCC. However, several recurrent sites of insertion in key genes associated with carcinogenesis have been described, these include *TERT* (telomerase reverse transcriptase), *MLL4* ( mixed-lineage leukaemia 4), *CCNE1* (cyclin E1) , *CCNA2* (cyclin A2)and *RARB* (Retinoic Acid Receptor Beta)[160,162]. The insertion of HBV DNA frequently occurs at promoter regions and results in overexpression, as has been demonstrated in *TERT* and *CCNE1*[160]. Chromosomal instability may also result from viral DNA integration. Wang *et al* described another integration site of viral DNA that resulted in the identification of the human cyclin A gene[163]. This integration resulted in an HBV pre-S2/S-cyclin A fusion protein with increased stability. Constitutive expression of this stabilized cyclin A protein may have led, or contributed, to increased cell proliferation. This effect was to cause HCC in the absence of underlying cirrhosis, further supporting the direct carcinogenic effect of HBV.

Another carcinogenic effect arises from the oncogenic viral proteins, e.g. HBx or the pre S2/S protein. The impact of these proteins has been suggested to be a key mechanism promoting carcinogenesis in humans through activation of several signalling pathways, control of apoptosis and DNA repair. In addition, frequent truncated HBx proteins are produced, and this protein could

increase metastasis or cell invasiveness *in vitro* models[164,165]. Other studies have suggested a link between the HBx protein and regulation of the mitotic spindle, resulting in chromosomal instability[166].

### 1.6.2 Pathogenesis of HCC relating to HCV

Hepatitis C is a single-stranded RNA virus belonging to the flaviridae virus family[155]. Unlike HBV, HCV cannot integrate into the host DNA, with the majority of its cell cycle occurring in the cytoplasm. Consequently, the development of HCC in the context of HCV results from the underlying cirrhosis, with a similar pathogenesis to alcohol and NASH. Experimental evidence suggests that HCV might operate through specific pathways that promote the malignant transformation of hepatocytes. These have included descriptions of modifications of signalling pathways like Wnt/B-catenine, TGF Beta, NFKappaB and P53, with these pathways having also been recognised as abberant in colorectal cancers[167,168]. Viral proteins, like those in HBV, have also been shown to have a carcinogenic role, namely HCV protein core or NS3, NS4B and NS5A[155,157]. Interpretation of these finding is difficult as the models used may not be transferrable to humans, particularly as most of these studies were performed in heterologous overexpression systems. With new and highly effective treatments for HCV answering this question may no longer be relevant.

### 1.6.3 Pathogenesis of HCC in alcoholic liver disease

Alcohol is a carcinogen and increases the risk of a number of cancers, including oesophageal, gastric, breast and colonic. Alcohol alone is likely to cause an increase in HCC, however, the vast majority of HCCs that occur in the context of excessive alcohol arise on a background of cirrhosis. Little is known about alcohol and non-cirrhotic HCC formation. In patients with HBV, HCV, NASH and haemochromatosis, excessive alcohol is known to significantly increase the risk of HCC when compared to a single factor[169,170]. Interestingly, patients with alcohol-related cirrhosis who stop drinking have a higher risk of HCC development than patients with persistent consumption[171]. This is likely to result from premature death in those who continue to drink, before the HCC has developed.

In the context, it is very difficult to separate alcohol consumption itself and underlying cirrhosis as two independent risk factors for HCC. When rodents have been given ethanol there is an increase in tumours, however this finding is inconsistent and often found in the presence of other known carcinogens[157,172]. Acetaldehyde is an ethanol metabolite that has carcinogenic properties through DNA binding. In addition, CYP2E1, a cytochrome induced by ethanol consumption, converts ethanol into acetaldehyde, but also into reactive oxygen species (like superoxide anion and hydrogen peroxide). Chronic oxidative stress induced by alcohol intake and cytokine production in

the context of chronic inflammation is a well-known phenomenon leading to cirrhosis and HCC development[157,173,174]. A further impact of excessive alcohol is an iron overloaded state, this is most noticeable in people with haemochromatosis who drink excessively, with a massive increased of HCC. Additional mechanisms of alcohol-mediated carcinogenesis exist, including down-regulation of levels of retinoic acid and modulation of methylation by down-regulation of S-adenosyl L methionine (SAMe)[175–177].

**1.6.4 Pathogenesis of HCC in non alcoholic fatty liver disease and non alcoholic steatohepatitis**

Obesity and diabetes are also direct risk factors for the development of HCC. Again the majority arise from a background of NASH and the subsequent chronic inflammation and cirrhosis, however a number of groups have described the occurrence of HCC in NASH without underlying cirrhosis[178,179]. This suggests a specific carcinogenic pathway outside the classic inflammation/cirrhosis pathways, but much like HCC in alcohol these are difficult to separate. The association between obesity and HCC appears to be derived from alteration in levels of cytokines and other inflammatory mediators, these include IL-6 TNF-alpha, leptin, adipokine and adiponectin. An abnormal balance between leptin and adipokine characterizes NAFLD and predisposes to HCC development. A decreased level of adiponectin, an anti-inflammatory, anti-proliferative cytokine exclusively produced by adipocytes, and a parallel increase in leptin, a pro-inflammatory, angiogenic and profibrogenic cytokine predominantly produced by adipocytes, are cardinal features of NAFLD and appear to increase the risk of HCC[157,180,181].

NASH is associated with lipotoxicity, secondary to lipid droplet deposition in the liver, which has been shown to directly contribute to HCC formation. In addition, there is also insulin resistance, resulting in hyperinsulinaemia and excessive production of IGF1, along with other growth factors such as insulin-like growth factor II, insulin receptor substrate 1, transforming growth factor-a and b as well as hepatocyte growth factor (HGF)[155,182]. These abnormal levels of growth factors alter signalling pathways that result in promotion of cell proliferation and inhibition of apoptosis.

From the various animal models described it would appear there are specific carcinogenic pathways separate from those associated with cirrhosis, but these require validation in human studies before they can be fully understood.

## 1.7 Genetic alteration associated with HCC formation

### 1.7.1 Oncogenes and tumour suppressor genes

Activation of pro-oncogenes, particularly of the *ras* family have been described in spontaneous and chemically induced rodent hepatocarcinogenic models[155]. In human HCC studies high levels of cyclin-D have been described, thus promoting cellular proliferation via *ras* signalling, this however, has only been described in up to 20% of cases, making its significance questionable[183].

The somatic genetic alterations first identified in HCC were activating mutations of beta-catenin (CTNNB1) and inactivating mutations of *p53*, in 20-40% and 20-50% of cases, respectively. The alteration in *p53* is significantly overexpressed in HCC associated with HBV[184,185]. Aflatoxin B1 is a well-recognised carcinogen for the formation of HCC. It is a contaminant of foods that are commonly consumed in Asia and Africa, for example chilli peppers, corn, cotton seed, millet, peanuts, rice, sorghum, sunflower seeds, tree nuts, wheat, and a variety of spices, with a linear relationship between levels of ingestion and incidence of HCC. Aflatoxin has been shown to cause specific mutations in the *p53* gene (R249S), as a result of its specificity this mutation can be used a biomarker of aflatoxin B1 exposure in patients with HCC[155,185]. In areas were aflotoxin levels are high this mutation is seen in up to 70% of HCCs, whilst being present in <5% were aflatoxin levels are low[186].

Activation of telomerase, preventing the shortening of telomeres, is frequently found in HCC. Recent evidence suggests that telomere dysfunction, leading to telomere based chromosomal instability, may be operative during the early stages of carcinogenesis while telomerase activation occurs during HCC progression[155,187].

Chronic inflammation and hepatic fibrosis underlie the carcinogenesis of hepatocytes. This results from increased cell turnover, oxidative DNA damage, activation of cellular oncogenes, the inactivation of tumour suppressor genes DNA mismatch repair defects and impaired chromosomal segregation, over expression of growth and angiogenic factors, and telomerase activation. Despite this, there are clear aetiology-specific processes that contribute to the formation of HCC and modify the disease course.

## 1.8 Biomarkers for colorectal cancer and HCC

In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic response to a therapeutic intervention[188].”

The World Health Organisation (WHO) has sought to define biomarkers further as “any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease[189]”. Using this board definition of a biomarker, clinical signs such as height and weight could be considered as biomarkers, however, the term is now typically shorthand for “molecular biomarker.” Molecular biomarkers can themselves take many forms, and as a consequence there are many strategies available for their discovery and validation. When considering malignancy any measurable specific molecular alteration of a cancer cell either at the DNA, RNA, protein, or metabolite level can be referred to as a cancer biomarker. An ideal biomarker for cancer would have applications in determining predisposition, early detection/screening, assessment of prognosis, and predicting drug response or amenability to therapy. No one single biomarker could meet all these needs and so combinations of specific applications of biomarkers are often performed.

There has been a great deal of work recently to find and define the role for non-invasive biomarkers, particularly for cancer. In order to prevent disorganized and uncoordinated biomarker research, Cancer Research UK’s Biomarker Discovery and Development Committee has suggested ‘roadmaps’(figure 6), which define a succinct research pathway. These are broadly divided into four chronological sections[190].

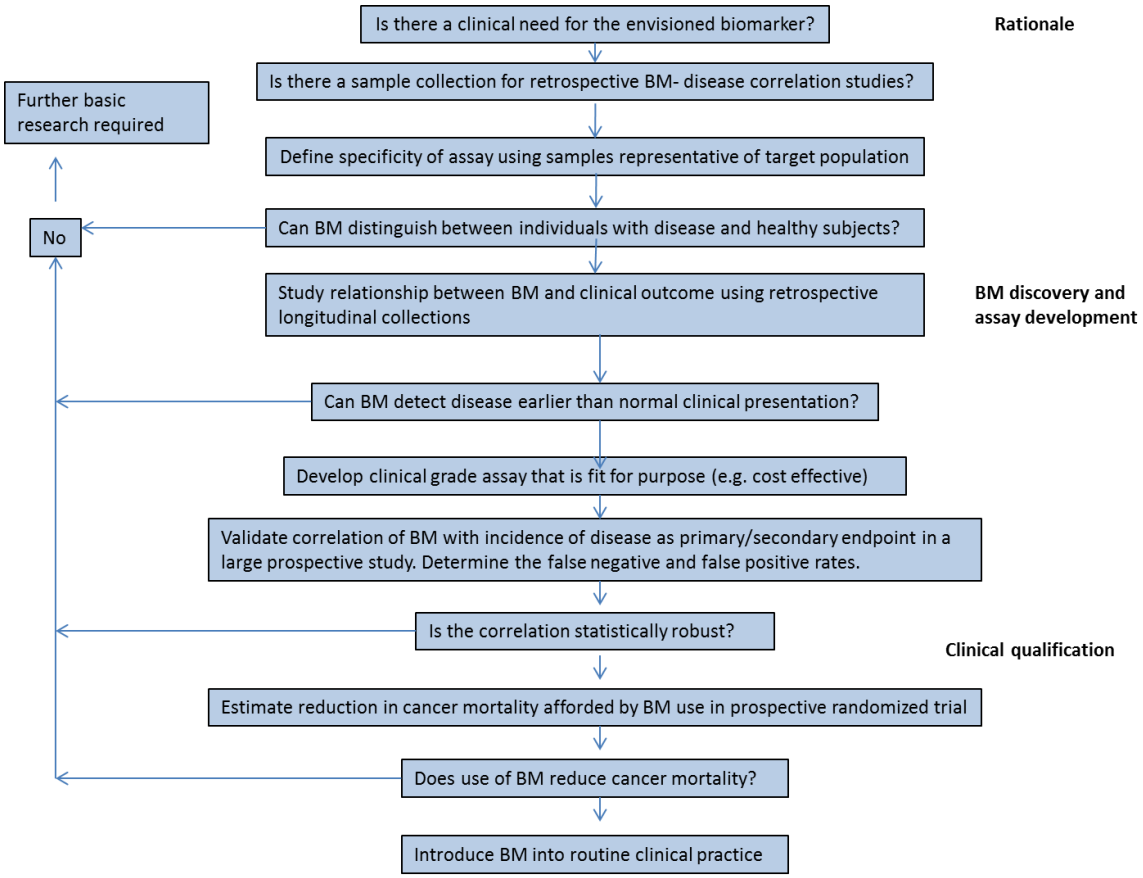
1) Rationale: does the envisioned biomarker address an unmet clinical need?

2) Biomarker assay development: the assay must be accurate and reproducible. The assay should be simple and cost effective, and performed to good clinical laboratory practice standards (as laid out by the British Association of Research Quality Assurance).

3) Biomarker discovery: the distribution of the biomarker in an appropriate sample population should be defined, and a retrospective analysis of the relationship between the biomarker and clinical outcome should be performed.

4) Biomarker clinical qualification: the relationship between the biomarker and clinical outcome should initially be assessed in a large retrospective analysis. Then, a large, prospective randomized study should be performed to assess the impact of the biomarker on clinical outcome.

**Figure 5: Screening biomarker roadmap.**




Adapted from Newton *et al*[190]

A key feature of biomarkers, particularly those used in colorectal cancer, is the ability to detect early stages of the disease or a pre-cancerous stage e.g. adenomatous polyps. This in turn allows the biomarker to have applications in screening. Colonoscopy is a key part of colorectal cancer screening, as it allows for confirmation of diagnosis, definition of morphology, site and histological analysis. As discussed, FOBt is the most widely used test, but this has low sensitivity and specificity, leading to a high normal colonoscopy rate. Ideally the biomarker used to indicate the need for colonoscopy, should have high sensitivity and specificity, while producing a low number of false-negative and false-positive results, to prevent subjecting healthy individuals to unnecessary colonoscopies. In addition to such detection biomarkers, prognostic markers which can predict the probable course of the cancer, stratification markers which can predict the response to drugs prior to beginning treatment, and efficacy markers which can monitor the efficacy of drugs treatment may also reduce the mortality rate of colorectal cancer[191]. Biomarkers can also be defined by the site of acquisition e.g. stool, serum, plasma and tissue.

### 1.8.1 Colorectal cancer biomarkers

FOBt has been discussed earlier in this chapter.

#### 1.8.1.1 Faecal DNA and genetic markers

Colonocytes are continuously shed from the mucosa into the faecal stream, cellular shedding from colorectal cancers occurs at an even greater rate[101]. The apoptotic resistance of cancerous cells means there is greater potential to detect the intact genomic DNA (L-DNA) for analysis, along with MSI detection. These cells found within stool can then be analysed for known genetic targets, such as, K-*ras*, *p53*, and adenomatous polyposis coli (*APC*)[191].

A systematic review conducted in 2014 considered seven studies investigating faecal DNA markers, looking at DNA hypermethylation of a single gene, or of a panel of genes. Overall sensitivities for colorectal cancer detection by faecal DNA markers ranged from 53% to 87% with varying specificities all of which are greater than 76%. Adenoma detection sensitivity ranged from 17% to 61%[192]. Boynton *et al* amplified six genomic fragments of different length from each of four different genetic loci (*APC*, *p53*, *BRCA1*, and *BRCA2*) using faecal specimens collected from 25 colorectal cancer patients and 77 controls. The specificity for colorectal cancer detection was 97% and the sensitivity was 57%[193].

Other studies that have directly compared faecal DNA markers to FOBt have shown that the DNA analysis has greater sensitivity than the FOBt without reduced specificity. The same studies show improved detection, by the DNA marker system, of invasive colorectal cancer and adenomatous polyps with high grade dysplasia, when compared to FOBt, 40.8% versus 14% [94]. When a combination of faecal DNA analysis and FOBt is performed the highest sensitivities and specificities can be achieved[191].

#### 1.8.1.2 Pyruvate kinase isoenzyme type M2 (M2PK)

M2-PK is an isoenzyme of pyruvate kinase, a key enzyme within glycolysis which catalyzes the ATP-producing conversion of phosphoenolpyruvate (PEP) to pyruvate. Depending upon the metabolic functions of the tissues, different isoenzymes of pyruvate kinase are expressed[194]. During tumourigenesis other isoenzymes of pyruvate kinase are lost and there is increased expression of the M2 subtype[195]. M2 pyruvate kinase can exist in two distinct forms, these are a nearly inactive dimeric form and a highly active tetrameric form. In tumour cells, M2-PK is mainly found to be in the dimeric form and has therefore been termed “Tumour M2-PK”. This dimerisation is induced by interaction with a number of oncoproteins, making it specific to cancerous states[194]. M2-PK is

released by tumour cells into blood and specifically by colorectal cancers into the faecal stream. Quantification of M2-PK in stool can, therefore act as a biomarker for colorectal cancer.

A meta-analysis, in 2012, concluded the use of M2-PK in screening would "close a clinical gap", due to its high sensitivity and specificity. 704 patients with colorectal cancer and 11,412 healthy subjects, from seventeen independent studies, were included in this meta-analysis. The mean faecal M2-PK sensitivity was 80.3%; the specificity was 95.2%. Four studies compared faecal M2-PK head-to-head with guaiac-based faecal occult blood test (gFOBt). Faecal M2-PK demonstrated a sensitivity of 81.1%, whereas the gFOBt detected only 36.9% of the colorectal cancers[194]. A further meta-analysis in 2014 looked at six papers relating to faecal biomarkers which assessed the same enzyme M2-PK as a potential biomarker in colorectal cancer detection. These studies used a sandwich ELISA to measure M2-PK activity and reported overall sensitivities ranging from 68% to 91%[192].

M2-PK has also been shown to have superiority in detecting adenomatous polyps, including non-bleeding lesions. Koss and colleagues showed the tumour M2-PK assay could be utilized to detect adenomas with a sensitivity of 60%[196]. Eight studies were described in the meta-analysis from 2012, this included 554 patients, with the following sensitivities: adenoma < 1 cm in diameter: 25%; adenoma > 1 cm: 44%; adenoma of unspecified diameter: 51%. In a direct comparison with gFOBt of adenoma > 1 cm in diameter, 47% tested positive with the faecal M2-PK test, whereas the gFOBt detected only 27%[194].

### 1.8.2 Serum Markers

#### 1.8.2.1 Carcinoembryonic antigen (CEA)

Carcinoembryonic antigen (CEA) is a high molecular weight glycoprotein belonging to the immunoglobulin superfamily. It is involved in cell adhesion, being produced during normal foetal development, however after birth its production ceases[190,191]. CEA is one of the most widely studied biomarkers in colorectal cancer[197]. It has been shown to be elevated not only in colorectal cancer but also in a number of other cancers including pancreatic, gastric, lung and breast, as well as several non-cancerous conditions including diabetes mellitus, ulcerative colitis, pancreatitis and heavy smokers.[190]. A number of recent meta-analyses have concluded that CEA has no role as a screening or diagnostic tool, particularly in an asymptomatic patient, and are in line with the American Society of Clinical Oncology (ASCO) and European Group on Tumour Markers (EGTM) recommendations[190,191,197]. This conclusion stems from the lack of sensitivity and specificity provided by CEA. The summary estimates from the meta-analysis performed by Liu *et al*

found the sensitivity and specificity for CEA in colorectal cancer to be 46.1% (95% CI: 44.8–47.4%) and 89.2% (95% CI: 88.2–90.2%), respectively.

The role for CEA lies in prognostics, directing treatment and detecting recurrence. A number of studies have demonstrated the use of CEA as a prognostic indicator, both as an independent factor and in combination with other factors[198–200]. Colorectal cancer patients with higher CEA levels have poorer prognosis[191]. Baseline CEA levels have been shown to be an independent predictor of treatment response, to bevacizumab-based treatment in metastatic disease. Baseline CEA serum levels inversely correlated with therapeutic response in patients receiving bevacizumab-based treatment (disease control rate, 84% vs 60%)[201]. Elevated CEA in patients undergoing hepatectomy with curative intent has been shown to indicate recurrence with high accuracy. Hara *et al* showed the post-test probabilities of recurrence in post-hepatectomy colorectal cancer patients with positive and negative serum CEA were approximately 70-90% and 10%, respectively[202].

#### 1.8.2.2 Carbohydrate antigen (CA19-9)

Carbohydrate antigen (CA) 19-9 is the second most investigated tumour marker relating to colorectal cancer, after CEA. CA 19-9 was originally defined by a monoclonal antibody produced by hybridoma prepared from the spleen cells of mice immunized with the human colorectal cancer cell line, SW 1116[191]. It is the most widely used biomarker for adenocarcinoma of the pancreas. However, this does not translate into use for colorectal cancer, due to its inferior sensitivity when compared to CEA, with both ASCO and EGTM not recommending its use in colorectal cancer[197,203].

#### 1.8.2.3 Tissue inhibitor of metalloproteinase type (TIMP)-1

Tissue inhibitor of metalloproteinase type (TIMP)-1 is a multifunctional glycoprotein which inhibits most matrix metalloproteinases (MMPs)[191]. Studies have shown TIMP-1 to be elevated in the serum of patients with colorectal cancer when compared to healthy volunteers[204]. It has also been shown not to be significantly elevated in patients with colonic adenomas and IBD[205]. The use of TIMP-1 appears to be limited to predicting prognosis, as it appears elevation of TIMP-1 is restricted to advanced colorectal cancer. Holten-Andersen *et al* demonstrated this relationship, their study showed a highly significant association between pre-operative plasma TIMP-1 levels and survival in colorectal cancer patients, with higher plasma TIMP-1 levels being associated with a poor outcome. Independent of clinical parameters including Dukes' stage, plasma TIMP-1 levels were found to strongly predict prognosis of colorectal cancer patients[206].

#### 1.8.2.4 Five-Serum-Marker panel *(Spondin-2, DcR3, Trail-R2, Reg IV, MIC 1)*

A study looking at the serum of 600 patients taken from patients with colorectal cancer, normal controls and benign disease evaluated four serum biomarkers, spondin- 2, tumour necrosis factor receptor superfamily member 6B (DcR3), TRAIL receptor 2 (TRAIL-R2) and Reg IV. All four markers, as well as a fifth marker, macrophage inhibitory cytokine 1 (MIC1), were found to be elevated in patients with colorectal cancer when compared to normal controls and patients with benign disease. The study also demonstrated a better sensitivity and specificity than CEA, suggesting a potential to improve early diagnosis[191,207].

#### 1.8.2.5 Collapsin response mediator protein-2 (CRMP-2)

Wu *et al* identified collapsin response mediator protein-2 (CRMP-2) and evaluated it as a potential colorectal cancer biomarker in the serums of 201 colorectal cancer patients and 210 healthy controls (48). The use of CRMP-2 alone showed better sensitivity but poorer specificity than CEA.

However, combined detection using CEA and CRMP-2 produced better sensitivity (77%) and specificity (95%) than detection using either of these markers alone. CRMP-2 may therefore, be a valuable serum marker when used in combination with CEA[191,207,208].

#### 1.8.2.6 MicroRNAs as biomarkers

The Tenth Annual MicroRNA (miRNA) as Biomarkers and Diagnostics Conference, 2014, was primarily focused on recent advancements in the field of miRNA in the early detection of disease, monitoring tumour growth/progression and its potential for precision medicine. The predominant outcomes from this meeting were that miRNAs are now well-established as regulators of tumourigenesis and can be utilized not only as potential biomarkers for the diagnosis and prognosis of a disease but also are useful in patient stratification and treatment response[209].

A number of miRNAs have been identified in colorectal cancers, both with up and down regulation, for example miR-34a[191]. These have been shown to interact with key oncogenes and tumour suppressor genes, such as APC and *p53*[210]. The significance of miRNAs in predicting treatment response with 5-fluorouracil based chemotherapy was described by Ju *et al* in 2013. Several miRNAs (miR-192, miR-215, miR-140, miR-129, let-7, miR-181b, miR-200 s) were found to be associated with chemoresistance by regulating key cell death pathways such as apoptosis and autophagy[211]. Other such important miRNAs have also recently been described, for example miRNA-215. miR-215 was identified to suppress the expression of both thymidylate synthase and dihydrofolate reductase. In addition, the expression of miR-215 was directly regulated by p53. The expression of miR-215 was significantly associated with colorectal cancer patient survival[209,212]. Other miRNAs that have

been shown to influence prognosis are hsa-miR-21 and hsa-miR-106a[213,214], with hsa defining the human origin of the miRNA.

It can be concluded therefore, that miRNAs appear to have a role in predicting survival and response to potential chemotherapy treatments.

### 1.8.3 Genetic and DNA markers in tissue

Detection of genetic alteration in tissue either acquired from biopsy or following resection has a potential as a biomarker. These markers have a role in predicting response or appropriateness of treatment and prognosis.

#### 1.8.3.1 *KRAS/B-RAF* and anti-EGFR therapy

Cetuximab and panitumumab are monoclonal antibodies that target EGFR, they have a role in the treatment of metastatic colorectal cancer. The abundance of EGFR in tumours has been shown to correlate poorly with response and is, therefore, an inadequate biomarker[215]. These drugs are effective in approximately 10% of unselected chemotherapy treatments. It has been shown, by many studies, that in chemotherapy-refractory metastatic colorectal cancer, a KRAS mutation predicts a complete lack of response to anti-EGFR therapy[79,190]. The phase III randomised trial of panitumumab showed almost a 2 fold increase in survival when the *KRAS* mutation was absent. Similar results have been noted in phase III trials of cetuximab[216,217]. Measurement of *KRAS* before commencing such treatment is recommended as routine practice by both the FDA and European Medicines Agency.

*B-RAF* is a mutant KRAS gene accounts for only 30–40% of the 90% of patients who do not respond to anti-EGFR therapy[190]. Nicolantino *et al* examined the effects of the *KRAS/B-RAF* mutation on treatment success with the anti-EGFR drugs. As fully expected those with the mutant *KRAS* gene all failed therapy. They also found that all those with the *B-RAF* mutation did not respond, whilst all the responders were free of the *B-RAF* mutation. Being positive for this mutation also conveyed a negative outcome for both progression free survival and overall mortality[218]. The evidence would appear to suggest that *B-RAF* testing should take place alongside the current practice of *KRAS* assessment.

### 1.8.4 Hepatocellular carcinoma biomarkers

The development of biomarkers in HCC has been thwarted by insufficient accuracy, making translation into routine clinical use difficult. Recently, advances in technology and an increased understanding of HCC biology have led to the discovery of novel biomarkers, which are currently in a

variety of development stages (Table 2 and 3). There are a number of serological markers for HCC, these include alpha-fetoprotein (AFP), MAGE-4 protein, des-gamma-carboxy prothrombin (DCP), the ratio of glycosylated AFP (L3 fraction) to total AFP, alpha-fucosidase, and glypican 3, transforming growth factor-beta (TGF-Beta) and most recently microRNAs (MiRNA). The application of a biomarker in HCC may be influenced by the underlying aetiology i.e. viral hepatitis, as they can be associated with specific genetic alterations and oncogenic protein formation. Despite this relatively large number of potential biomarkers very few are routinely used in clinical practice and, of those that are, their use is becoming more limited as knowledge and experience of their use grows.

**Table 2: Biomarkers for the early detection of HCC and their stage of development.**

| Biomarker | Phase of development |
|---|---|
| AFP | 5 |
| AFP-L3 | 2 |
| DCP | 2 |
| Osteopontin | 2 |
| GP73 | 2 |
| Glypican | 2 |
| SCCA | 2 |
| microRNA | 2 |
| Canavanionsuccinate | 2 |
| Glycochenodeoxycholic acid | 2 |
| Urine glycocholic acid | 2 |

**Table 3: Phases of biomarker development for early detection[218].**

| Phase of development | Study design | Description |
|---|---|---|
| 1 | Pre-clinical | Identification of potential candidates |
| 2 | Case control studies | Characterisation of diagnostic performance |
| 3 | Retrospective longitudinal | Ability to detect pre-clinical disease |
| 4 | Prospective screening | Define sensitivity and specificity |
| 5 | Randomised controlled | Determine ability to reduce mortality |

#### 1.8.4.1 Alpha-fetoprotein

AFP is the biomarker that has been most prominent over recent years but most recently has fallen out of favour due to its low sensitivity, particularly at a lower cut-off value[141]. It is a glycoprotein produced by the foetal liver and yolk sac during pregnancy. A systematic review evaluating AFP in cirrhotic patients, using a cut-off level of 20ng/mL, showed sensitivities and specificities of 41-65% and 80-94% respectively HCC at any stage[219]. Although 20ng/mL is the most commonly used cut-off in clinical practice, this value was derived from a study in which only one-third of patients had early HCC (70 HCC and 170 chronic liver disease)[220]. The detection of HCC based on a significant rise of AFP in the absence of a positive US is considered to be exceptional[140]. Furthermore, only a small proportion of early HCCs are associated with a rise in AFP serum levels, with studies clearly demonstrating the correlation between tumour size and serum AFP levels[143]. Approximately 40% of patients with early HCC and 15-20% of patients with advanced HCC show normal AFP levels. In China, where HCC is extensively studied, around 30-40% of HCC patients are AFP negative[142]. Tumours that are associated with increased AFP, should be considered to have a greater malignant potential and a poorer prognosis[142]. The role of AFP is now for prognostic estimation and treatment monitoring, rather than surveillance. Of the other potential biomarkers, most have been described in a diagnostic setting rather than surveillance. Some, namely, DCP and glycosylated AFP-total ratio, have been shown to be associated with advanced disease, including local invasion, thus making them redundant as a screening tool[221]. None has been shown to be superior to AFP alone, therefore they are not currently recommended for clinical use[129].

Recent studies have shown the variation of AFP when considering different aetiologies and levels of hepatic inflammation. These studies have suggested the accuracy of AFP may be improved by adjusting for degree of hepatic inflammation and/or aetiology of liver disease[222]. It would appear

that AFP tends to correlate with ALT levels in patients with chronic viral hepatitis in the absence of HCC, whereas AFP often increases disproportionately to ALT levels in those with HCC[223]. Also, recently described is the impaired accuracy and specificity of AFP in HCV-positive patients compared to negative patients, specificity 0.89 vs. 0.83, p = 0.007[224].

#### 1.8.4.2 AFP-L3

Lens culinaris agglutin-reactive AFP (AFP-L3) is the glycosylated isoform of AFP. It has been proposed as a biomarker for the early detection of HCC as it has a superior specificity compared with AFP. During its early development it was found to have a specificity approaching 90%[225]. Further assessment of AFP-L3 in large cohort studies, including the Early Detection Research Network (EDRN) multi-centre phase II study, support this finding with specificities above 90%. However, sensitivity across all these studies was low, ranging from 37-60%[226,227]. AFP-L3 is reported as the proportion of AFP-L3 to total AFP, with a cut-off of 10% traditionally used for HCC detection. The development of a highly sensitive assay for AFP-L3 (hs-AFP-L3) has improved sensitivity. It means that APF-L3 can be measured in those with lower AFP levels, in fact hs-AFP-L3 can be measured in patients with AFP as low as 2 ng/mL[222]. The sensitivity of hs-AFP-L3 for early stage HCC appears to be significantly higher than that of conventional AFP-L3 (49% vs 17%)[228]. Phase III biomarker studies are currently in progress to assess the utility of AFP-L3.

#### 1.8.4.3 Des-Gamma-Carboxy Prothrombin (DCP)

DCP is produced by malignant hepatocytes. It is an abnormal prothrombin protein that is generated as a result of an acquired defect in post-translational carboxylation[222]. DCP production is independent of vitamin K deficiency, although pharmacological doses of vitamin K can transiently suppress DCP production in some tumours[229]. Normalization of DCP levels correlates with successful tumour resection and appears to be a marker of tumour activity. Plasma DCP does not correlate with AFP levels. However, they do show similar traits in terms of the correlation with increasing tumour size. In tumours of less than 3 cm, DCP levels are increased in only 20% of patients[229,230]. An earlier phase II biomarker study reported DCP had a sensitivity and specificity of 89% and 95% for differentiating patients with HCC from those with cirrhosis and chronic hepatitis[231]. Further phase II studies and several prospective cohort studies failed to demonstrate this high level of sensitivity, with sensitivities ranging from 23 to 57% when DCP was used alone [232,233]. Secondary analysis from the HALT-C trial showed a significant improvement in the sensitivity for tumour detection from 58% to 87% three months prior to HCC presentation, suggesting that there is a role for DCP in early detection of HCC[234]. Phase III biomarker studies are currently in progress to further assess the utility of DCP.

#### 1.8.4.4 Glypican-3

Glypican-3 (GPC3), a membrane-anchored heparin sulfate proteoglycan, has been demonstrated to interact with growth factors and modulate their activities and, as such, it plays an important role in cell proliferation and survival regulation. It binds to the cell membrane through the glycosylphosphatidylinositol anchors. GPC3 mRNA was up-regulated significantly in tumour tissues of HCC compared to liver tissues of healthy adults, and liver tissues of patients with non-malignant liver disease. It can be detected in 40–53% of HCC patients and 33% of HCC patients seronegative for both AFP and DCP[229,235,236]. GPC3 demonstrated 100% specificity among 50 patients with chronic hepatitis and 50 with intrahepatic cholangiocarcinoma and 98% specificity among 50 patients with cirrhosis[237]. Much like DCP, the good specificity is associated with an insufficient sensitivity, with a number of studies showing it to be consistently around 50%, but this has been shown to been independent of tumour size, allowing early tumour detection[222]. GPC3 positivity has also been shown to be an independent negative prognostic indicator in patients with HCC[238].

#### 1.8.4.5 Osteopontin

Osteopontin is an integrin-binding phosphoprotein that is over-expressed in several cancers including lung, breast and colon cancer[222]. One of the first descriptions of osteopontin was in 2006, when 62 patients with HCC were compared to 60 with cirrhosis only and 60 healthy controls. This study demonstrated a sensitivity and specificity of 87% and 82%[239]. Further assessment of osteopontin, in a Thai cohort, demonstrated a 93% sensitivity and 61% specificity. Attempts have been to improve this by combining measurement with AFP levels. In a US this cohort: the combination of AFP and osteopontin had a sensitivity of 83% and specificity of 63% for early HCC[240]. Assessment of osteopontin, either alone or in combination with other biomarkers, is currently underway in phase III biomarker trials.

#### 1.8.4.6 Alpha-l-fucosidase (AFU)

Alpha-l-fucosidase (AFU) is a lysosomal enzyme found in all mammalian cells with a function to hydrolyze fucose glycosidic linkages of glycoprotein and glycolipids, its activity increasing in the serum of HCC patients[229]. There have been reports that a raised AFU level could act as a precursor to HCC development, with 85% of patients in a study having an elevated AFU at least 6 months before the detection of HCC by ultrasonography[241]. Sensitivity and specificity levels of 81% and 70%, respectively, have been reported with a cut-off of 870nmol/mL/h. This could potentially be improved further with the combined measurement of AFP and AFU[229].

#### 1.8.4.7 Golgi protein-73

Golgi Protein-73 (GP73) is a trans-membrane glycoprotein that is normally expressed in epithelial cells but whose biologic function is unknown[222]. Golgi-apparatus-associated protein has been shown to have a higher sensitivity than AFP in the detection of HCC[242]. This study found GP73 to have a sensitivity and specificity of 69% and 86% respectively; of note, GP73 levels were elevated in 32 of 52 HCC patients with an AFP < 20ng/mL. A study looking at serial ELISA measurements of GP73, in patients with HCV, appears to demonstrate that it is a promising complimentary serum marker in the surveillance of HCC[243]. Elevated GP73 levels have been shown to be independent of cirrhosis aetiology, liver function, and tumour size. When used in combination with AFP, sensitivity and specificity have been shown to improve to 98% and 85%, respectively. Systematic reviews of the literature associated with GP73 have drawn different conclusions about its clinical application: those that have criticised it have suggested that the studies have been of poor quality and conducted in non-representative cohorts, while others have suggested that GP73 is superior to AFP for the detection of early HCC, but concluded that further phase III biomarker studies are required.

#### 1.8.4.8 Transforming growth factor-beta (TGF-Beta)

Belonging to a superfamily of polypeptide signalling molecules involved in regulating cell growth, differentiation, angiogenesis, invasion, and immune function, TGF-beta is a predominant form of the growth factor family in humans[229]. TGF-Beta has been shown to be elevated in the serum of patients with HCC when compared to healthy controls and cirrhotic patients without HCC, along with the over-expression of mRNA[244,245].

#### 1.8.4.9 Squamous cellular carcinoma antigen (SCCA)

Squamous cellular carcinoma antigen (SCCA) is a serine protease inhibitor that is physiologically present in squamous epithelium[222]. Evaluation of SCCA, in 2005, assessed 210 patients with HCC and 90 with cirrhosis and found that it had a sensitivity of 84% but a specificity of 49%[246]. Further assessment by the same group looking specifically at early lesion detection (<3cm) demonstrated a sensitivity of 56%, and specificity of 75%[247]. Systematic reviews have concluded that SCCA is inferior to AFP and so it is unlikely to find a place in clinical practice.

#### 1.8.4.10 Micro RNA

MicroRNAs (miRNAs) are a family of endogenous, small (21–23 nucleotides), non-coding but functional RNAs, which have been found in worms, flies, and mammals including human beings. MicroRNAs regulate gene expression by binding to specific messenger RNAs and prevent their translation into protein. They can act to down regulate a large number of genes at any one time, which in turn can have profound effects on cellular programmes, including repair and death[248].

As a result, some miRNAs are involved in direct HCC carcinogenesis by promoting cancer stem cells and by controlling cell proliferation and apoptosis. Others are associated with HCC progression by influencing cell migration and invasion[229]. miR-124 and miR-224 are greatly up-regulated in human HCC when compared to both paired pre-cancerous cirrhotic tissues and cirrhotic livers without HCC, and play a role in cell proliferation, migration, invasion and anti-apoptosis[249]. miR-101 is a tumour-suppressive miRNA. It is significantly under-expressed in multiple types of cancers including HCC[250]. The role of miRNAs as biomarkers for HCC has yet to be clearly defined but they appear to have a place in diagnosis and prognosis.

Because of the large number of potential biomarkers that have been described in this condition, selecting a single agent or combination that will provide suitable clinical applications is difficult.

## 1.9 Human gut microbiome, colorectal cancer and volatile organic compounds

### 1.9.1 The human gut microbiome

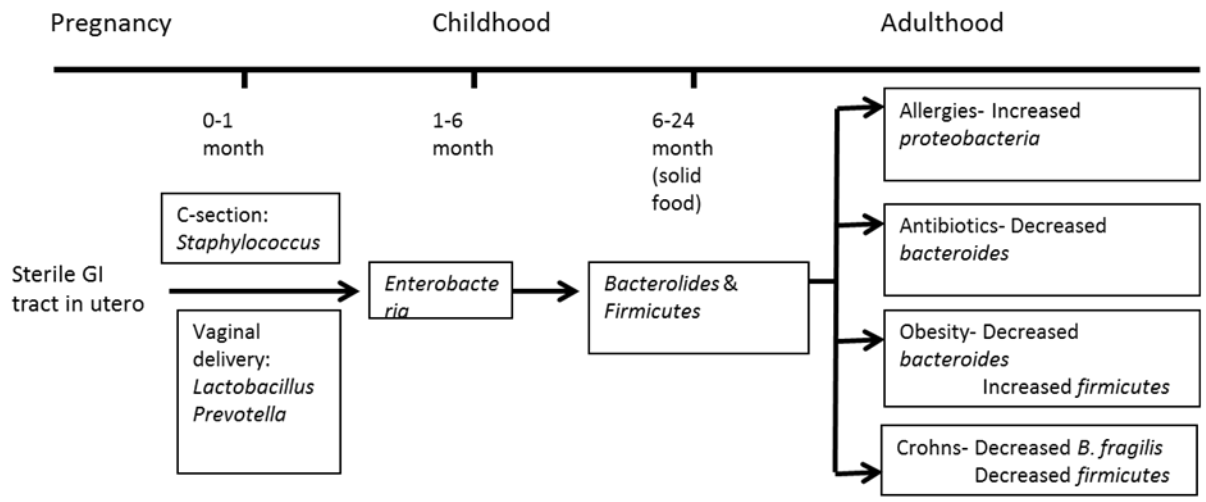
The human body consists of far more than just human cells, it plays host to a vast number of microbial cells, including bacteria, viruses, eukaryotes and fungi. Collectively, the microbial organisms that reside in and on the human body constitute our microbiota, and the genes they encode are known as the microbiome. Advancements in culture-independent high-throughput sequencing has increased our knowledge of the microbiota species diversity[251,252]. In doing so it is said that the human gut contains some $10^{13}$ different bacteria[253], only a small minority of these can be cultured. With an individual human colon containing approximately 160 species[254].

Following birth, as we develop from infancy into adulthood, so too does our microbiota. It was originally thought that the foetal gut was sterile and free of microbes, therefore was essentially a blank canvas upon delivery. However, recent work has provided evidence that there are bacteria present in the amniotic fluid in utero in healthy foetuses and in neonates. although, the number and diversity is very low[255]. The first stool passed by an infant is known as the meconium, it has been shown to be free of viral matter but contains a small number of bacterial species, again suggesting the foetal gut may not be sterile as previously thought[256]. However, there is also clear evidence that the mode of delivery impacts on the bacteria that can be identified within the infant gut. Those born vaginally have been shown to have a gut microbiota that resembles that found in the vaginal microbiota of their mothers. Whilst those delivered via caesarean section demonstrate a microbiota characteristic of skin and dominated by taxa such as *Staphylococcus* and *Propionibacterium*

spp[257]. This suggests any bacteria detected in the meconium actually colonise the gut after exposure during the delivery process, rather than in utero.

This initially minimally diverse microbiota quickly increases shortly after birth, with both bacteria and viruses increasing in cell and species number[256,258]. Early colonizers are generally aerotolerant, as the gut initially contains oxygen, and then are replaced by anaerobes that are typical of the adult gut microbiota. These initial colonizers are replaced by species, that by 11 months, are specific to the infant and different from those found in the mother[251,258]. As time passes and the infant develops the composition and diversity of the microbiota changes (Figure 6). This change has been shown to be gradual process, but can be punctuated by significant changes associated with life events such as infections, use of antibiotics, environmental exposures and the introduction of solid foods. The bacterial composition begins to converge towards an adult-like microbiota by the end of the first year of life and fully resembles the adult microbiota by 2.5 years of age[256,259].

**Figure 6: Development of the human microbiota from foetal to childhood, then through to adulthood**




Dominant bacterial species shown, along with regulation of functionality[251].

Functional genetic analysis of the infant microbiome demonstrates a functionality found in mothers, despite having a different taxonomic composition, with some specific differences such as, an enrichment in genes to facilitate lactate utilization when the infant's diet is breast or formula milk. Other analyses have shown that the ability the utilise plant-derived glycans is present before the introduction of solid foods, suggesting that the infant gut is ready to switch to a diet not exclusively based on milk before the actual change in diet takes place[256,258].

Once the microbiota has reached maturity it remains mostly stable until old age[251]. The composition of the intestinal microbiota in older people (>65 years) is extremely variable between individuals, and differs from the core microbiota and diversity levels of younger adults[260]. A feature of the ageing process is immunosenescence, this term describes the negative impact that the ageing process has on the immune system, including its development and its ability to function. Alterations to the innate and adaptive immune systems are recognised as a feature of immunosenescence, the interaction of both the innate and adaptive immune systems is key for the gut microbiota health and function[261,262]. Whether the alteration in an ageing immune system results in alteration in the composition of the gut microbiota, or the host's ability to interact with an already altering diversity is not clear. Deterioration in dentition, salivary function, digestion, intestinal transit time and changes in diet may affect the intestinal microbiota upon ageing. Exposure to antibiotics and other medications that may occur through a person's life also affects the composition[251,263]. Antibiotics have a profound effect on the microbiota, with good evidence to suggest major alterations of the microbiota following treatment with antibiotics. As the established gut bacterial community reshapes after treatment with antibiotics, there is a reduced resistance to colonization, allowing foreign microbes that can outgrow commensal bacteria to cause permanent changes in the structure of the microbiota and varying states of disease[251,264].

### 1.9.2 Gut microbiota and colorectal cancer

As the understanding of the composition of the gut microbiota has developed, so too has the knowledge of its diverse function and influence of health, including in the development of colorectal cancer. The intestinal bacteria are essential in digestion and absorption of non-digestible carbohydrates (fibre), production of vitamins B and K, metabolism of endogenous and exogenous compounds, immune potentiation, are actively involved in innate and cell mediated immunity, help to maintain intestinal barrier function, help to maintain intestinal barrier function and assist with an appropriate immune response against pathogenic microbes[265]. This symbiotic relationship, normobiosis, is essential for maintaining overall gut and human health. When there is a shift in this delicate balance, dysbiosis, there can be detrimental effects for the host, including the development of colonic polyps and colorectal cancers[251,253,266,267].

The relatively long duration over which colorectal cancer develops and the multi-factorial nature of its aetiology have made it difficult to accurately define the exact role of microbiota. However, potential mechanisms have been described:

- Inflammation: it is becoming increasing clear that the microbiota has a major influence on immune response and induction of chronic inflammation[268].
- Altered metabolic profiles: either through reduction production or increased consumption of protective metabolites. Together with production of toxic/harmful metabolic by-products from dietary carbohydrates, protein, bile and mutagenic precursors e.g. phenol compounds[265,269]
- Cell wall antigens and bacterial colicins: *Streptococcus gallolyticus* (formally knonw as *Streptococcus* bovi)s antigens have been shown to promote premalignant lesions through aberrant crypt formation. *E.coli* strains have been shown to induce DNA double strand breaks in intestinal cells and trigger chromosomal instability, gene mutations and cell transformations[270,271].

Studies of immigrants show that cancer rates quickly match those of the current country of residence, even when they originate from a country with a lower incidence. Genes do not change over such a short period of time, therefore dietary changes and environmental changes that lead to dysbiosis are thought to be important contributors to the development of colorectal cancer[272]. A summary of recent human studies of gut bacteria associated with adenoma and colorectal cancer formation, is seen in Table 4:

**Table 4: Human studies of gut microbiota and metabolites associated with colorectal cancer.**

| **Study (author, year)** | **Sample material** | **Disease** | **Summary of findings** |
|---|---|---|---|
| Nugent *et al*. 2014 | Colonic tissue | Adenoma | Bacterial dysbiosis, with alterations specific to adenomatous tissue when compared to non adenomatous |
| Brim *et al*, 2013 | Faeces | Adenoma | Altered microbiome to subgenus level, overall genome existed unaltered |
| Chen *et al*. 2013 | Faeces | Adenoma | Lower SCFA production in adenoma group, specifically reduced butyrate producing bacteria than in healthy controls |
| Shen *et al*. 2010 | Colonic tissue | Adenoma | Higher relative abundance of *Proteobacteria* and lower abundance of *Bacteroides* |
| Zackular *et al* 2014 | Faeces | Adenoma and colorectal cancer | Higher relative abundance of *Fusobacterium, Porphyromonas* and *Enterobacteriaceae*, with lower *Clostridium* in adenoma and |

| | | | |
|---|---|---|---|
| | | | cancer |
| Ohigashi *et al* 2013 | Faeces | Adenoma and colorectal cancer | Decreased SCFAs and increased pH in colorectal cancer |
| Marchesi *et al*. 2011 | Colonic tissue | Colorectal cancer | High abudance of *Fusobacterium* in colorectal cancers compared to normal tissue |

Adapted from [272].

### 1.9.3 The microbiota and inflammation

As the colonic mucosa is constantly exposed to the gut microbiota and its metabolites, bacterial stimulation of immune responses has the potential to cause continuous low-grade inflammation. The innate and adaptive immune systems work to maintain a homeostatic state in the context of this low-grade inflammation. This homeostasis creates a balance between the pro-inflammatory mediators (e.g. IL-1B, IFNγ, IL-8, TNF-$\alpha$, IL-23, IL-12 IL-17 and IL-6) and the anti-inflammatory mediators (e.g. IL-10 and TGF-Beta). Disruption of this pro and anti-inflammatory state can lead to uncontrolled chronic inflammation and subsequently colorectal cancer. The ratio of the harmful and protective commensal bacteria is key to maintaining homeostasis and gut health. Variation in numbers and diversity contributes to increased mucosal permeability, bacterial translocation and increased activation of components of the innate and adaptive immune system to promote chronic inflammation[265,273].

Assessment of the microenvironment of colorectal tumours demonstrates several different immune cells types, the origin of which is felt to arise from alteration in the microbiota. The most numerous of these is tissue-associated macrophages (TAMs)[268]. TAMs have been shown to promote tumour growth, with high numbers of TAMs correlating with progression[274]. The next most abundant immune cells in the microenvironment are T cells, they have been shown to exert both a suppressing and promoting effect. $T_H1$ cells have a direct lytic impact on cancer cells and secrete cytotoxic cytokines that limit cancer progression, whilst IFN-γ producing T cells are associated with tumourigenesis. Inflammation in the absence of microbiota or its products seems to be insufficient to induce colorectal cancer[270]. This interaction is not only seen with sporadic cases but also in inherited conditions. Germ free mice with the *APC* gene have been shown to exhibit a two-fold reduction in the number of intestinal adenomas, when compared to *APC* mice with a conventional

microbiota[275]. Mouse models using chemically induced colorectal cancer models have also shown the importance of the microbiota. In this model antibiotics were used to "knock out" the microbiota, those mice that received antibiotics had significantly fewer tumours in the colon than untreated mice (p=<0.001)[266].

Another major consequence of the chronic inflammation induced by the dysbiosis of the microbiota is the activation of key pro-survival and pro-proliferative signalling pathways by transcription factors such as NF-κB and STAT3 in epithelial cells well as generation of reactive oxygen or nitrogen species that leads to oxidative stress, DNA damage, aberrant proliferation and, ultimately, development of colorectal adenomas and cancer[265,276].

### 1.9.4 The microbiota, metabolites and diet

The principal elements of the human diet, namely carbohydrates, protein and fat, that are undigested and reach the colon undergo bacterial digestion and fermentation by the anaerobic microbial community. This produces a widely diverse set of metabolites with both protective and potentially toxic effects. As such, the link between diet and colorectal cancer may be explained, in part, by the activities of the intestinal microbiota. The major fermentation products in healthy adults are gases and organic acids, particularly the three short-chain fatty acids (SCFAs) acetate, propionate and butanoate typically in a 3/1/1 ratio[268].

#### 1.9.4.1 Impact of SCFA

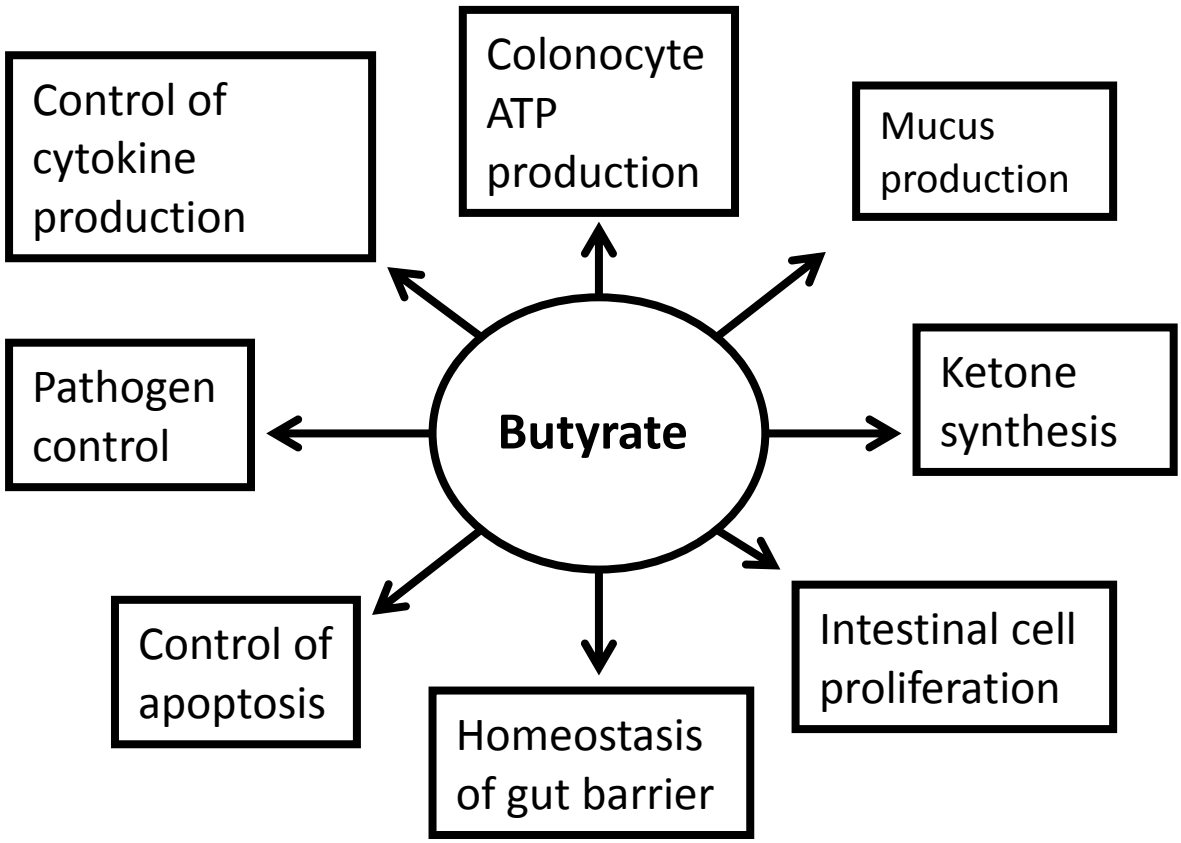
The production of SCFAs is influenced by diet, specifically the ingestion of non-digestable carbohydrates, such as resistant starches, pectins, gums, and cellulose. Higher levels of ingestion lead to increased colonic fermentation, increased gut transit, increased stool output and a decrease in the luminal pH. Different groups of enteric bacteria appear to favour the production of specific SCFAs. Acetate, propionate and butyrate are rapidly absorbed from the gut lumen, but their subsequent distribution, fate and effects on host cell metabolism differ. Butyrate (aka butanoate) is preferentially used as an energy source by gut epithelial cells, and its concentration in the systemic circulation is low. Propionate is mostly metabolized in the liver, and only acetate achieves relatively high concentrations in peripheral blood[268]. Approximately 95% of the butyrate produced by colonic bacteria is transported across the epithelium, but concentrations in portal blood are usually undetectable as a result of rapid utilisation as the preferential energy source of colonocytes[277]. There is an inverse relationship between butyrate concentrations and pH. More acidic conditions favour further fermentation and reduce the absorption of potential carcinogenic substrates. There is also an increased excretion of ammonia[278].

The luminal concentration of SCFAs, specifically butyrate, appears to be the key factor when considering their impact on colorectal cancer. In higher concentrations butyrate appears to exhibit a protective effect. Butyrate influences gene expression through its non-competitive inhibition of histone deacetylases (HDACs), leading to hyperacetylation of chromatin[279]. This has multiple consequences for gene expression and cellular differentiation, including the down-regulation of pro-inflammatory cytokines, such as IL-6 and IL-12, and inhibition of activation of the transcription factor NF-kappa B[268,280]. A further mode by which butyrate affects inflammation is via the induction of differentiation of regulatory T cells that express the transcription factor FOXP3[281].

In normal cells butyrate acts as a survival factor by acting as a major source of energy, but it is an inducer of apoptosis in human colonic carcinoma cells[282]. Studies have shown that the influence of butyrate over apoptosis is lost when the early DNA damage of the adenoma –carcinoma pathway has occurred[283]. This impact appears to be concentration dependent, as low amounts of butyrate enhance cell proliferation while high amounts inhibit it. Butyrate also conveys influence over metastasis once the carcinoma has developed. It has been shown to alter the expression of both pro- and anti-metastatic genes, with a net suppression of pro-metastatic and up-regulation of anti-metastatic genes[284]. It has also been suggested that butyrate inhibits tumour induced angiogenesis through interaction with two angiogenesis-related proteins, vascular endothelial growth factor (VEGF) and hypoxia-inducible factor (HIF)-1a[285].

Butyrate also shows a protective effect against oxidative damage by modulation of genes associated with oxidative stress. The expression of human catalase, human cyclo-oxygenase 2 and human metallothionein 2A is enhanced, whereas that of glutathione reductase, PG-endoperoxide synthase 2 and superoxide dismutase 2 is lowered[283]. Other effects of butyrate studied in multiple colonic cancer cell lines include the enhancement of the activity of the detoxifying enzyme glutathione-S-transferase[269].

**Figure 7: Multiple intestinal effects of butyrate.**




Adapted from Guilloteau *et al*[283].

Weir *et al* demonstrated how butyrate-producing species were under-represented in the analysis of stool from patients with colorectal cancer when compared to healthy controls, such as *Ruminococcus* spp. and *Pseudobutyrivibrio ruminis*, While a mucin-degrading species, *Akkermansia muciniphila*, was about 4-fold higher in colorectal cancer (p=0.01). Proportionately higher amounts of butyrate were seen in stool of healthy individuals while relative concentrations of acetate were higher in stools of colorectal cancer patients[286].

#### 1.9.4.2 Harmful metabolites

The metabolites that potentiate a harmful effect are derived primarily from protein fermentation in the colon. With increased protein ingestion, the colonic residue contains more sulphur, nitrates, ammonia, amines, branched chain amino acids, and hydrogen sulphide[265]. Some bacteria ferment aromatic amino acids to produce potentially bioactive compounds, including phenylacetic aicd, phenols, indoles and *p*-cresol. These compounds have been shown to exhibit carcinogenic effects and promote the development of colorectal cancer. They have been shown to be most abundant in

the distal colon where tumour development in most common[287]. These and nitrogen containing compounds, particularly *N*-nitroso, were shown to positively correlate with colorectal cancer in the European population. Ammonia, which is another product of protein fermentation, is also a potentially carcinogenic agent at relatively low concentrations, as shown by the increase in mucosal damage and colonic adenocarcinoma in a rat model[288,289].

Amine production, predominately in the form of putrescine, spermidine and spermine, are produced from arginine, both by human cells and gut bacteria. Certain bacteria are known to increase production by up-regulating synthesis by host cells e.g. *Bacteroides fragilis*. At physiological levels they serve an essential purposes, such as the maintenance of the structural integrity of membranes and nucleic acids, gene regulation and translation[290,291]. However, at high levels they have been shown to be toxic, inducing oxidative stress and potentially colorectal cancer[290].

Hydrogen sulphide is produced in the gut via the reduction of diet derived sulphate and through the metabolism of other compounds, e.g. sulphur amino acids[292]. Within the microbiota specific bacteria have been identified as specialist sulphate reducing organisms, for example *Desulfovibrio* spp. are present in most individuals, but have yet to be shown to be more numerous in patients with colorectal cancer[268,293]. Hydrogen sulphide exhibits a toxic effect through the inhibition of butyrate oxidation resulting in colonocyte barrier breakdown and DNA damage.

Diets that are high in fat are associated with increased bile secretion and thus increased concentration of bile acids in the colon. Increased faecal bile acid concentrations have been linked to colorectal cancer[294,295]. Animal models fed a high fat diet have been shown to have significant alteration in their microbiota and a fall in the production of SCFAs. Bile acids such as deoxycholic acid contribute to increased reactive oxygen species, DNA damage, genomic instability and tumour growth. Bacteria such as *Clostridium* spp. have been implicated in these processes, particularly reactive oxygen species and the subsequent chronic inflammation[268,296].

Consumers of a “western” diet have low fibre, high protein and high fat. Thus, they will have lower concentrations of SCFAs, higher levels of amino acid derived products and higher concentrations of bile acids. The interaction of these dietary compounds with the microbiota and the subsequent metabolism of the by-products can produce volatile organic compounds detectable in the faeces.

### 1.9.5 Carcinogenic impact of the microbiota

As previously described chronic inflammation of the colonic mucosa predisposes to the development of colorectal cancer. One example is the *Bacteroides fragilis*-mediated pathogenesis model of colorectal cancer[297]. In this model there are two types of *Bacteroides fragilis*, classified according

to their ability to secrete *B. fragilis* toxin. They are termed enterotoxigenic *B. fragilis* (ETBF) and nontoxigenic *B. fragilis* (NTBF). When mice are inoculated with ETBF they have been shown to develop colitis and colorectal cancer whilst those inoculated with NTBF develop neither[298].

*Fusobacterium* species, in particular *Fusobacterium nucleatum* have been implicated as having a direct carcinogenic effect in the development of colorectal cancer, with a significant over-representation in colorectal cancer tissue. *F. Nucleatum* is an invasive, adherent and pro-inflammatory anaerobic bacterium[299]. They have been shown to be more numerous in colonic tumour tissue when compared to normal colonic tissue[300]. This effect has been shown to be derived from the FadA adhesion protein complex (FadAc). Studies have demonstrated its interaction with E-cadherin and the Wnt signalling pathway, subsequently promoting invasion of the tumour cells and cell proliferation[301]. Castellarin *et al* also demonstrated a positive correlation with lymph node metastasis.

*E.coli* species have been demonstrated to have a carcinogenic role in colorectal cancer. Mice treated with azoxymethane, that subsequently develop colorectal cancer, have been shown to have an increased number of the NC101 strain of *E.coli*, that is known to promote highly malignant colonic tumours[270].

Transplantation of faeces from tumour bearing mice into germ free mice has been shown to result in an increase in inflammation and colorectal cancers in the recipient when compared to those not receiving transplanted stool. The same has been demonstrated with the transplantation of human stool from colorectal cancer patient into germ free mice[302,303].

Alterations in the gut microbiota that occur with age have also been linked to the development of colorectal cancer. As the host ages the total number of bacteria reduce and the composition changes with lower numbers of Firmicutes and increased proportion of *Bacteroidetes*[304]: a composition that has previously been shown to increase colorectal cancer. Others have also reported an increase in *Escherichia coli* and *Enterococci* spp. with increasing age, but this has failed to be consistently shown in similar studies[305].

### 1.9.6 Volatile organic compounds and colorectal cancer

Volatile organic compounds (VOCs) are a diverse group of carbon-based chemicals that are volatile at ambient temperature. They may be odorous and may be emitted from bodily fluids and, as a result, VOCs emitted from faeces and urine may include biomarkers of use in the assessment of GI and liver disease[306]. Their place as diagnostic tools can be traced back to Hippocrates who instructed his students to smell the breath of their patients, as well as pour human sputum on hot

coals, in order to produce a smell as a potential indicator of human diseases. The understanding of VOCs and their role in human disease has risen dramatically in the last two decades, with the introduction of solid phase microextraction (SPME) in 1989. This was combined with traditional techniques using a separation step with gas chromatography and analysis with mass spectrometry, electron capture detection or flame ionization detector. Gas chromatography and mass spectrometry in combination with SPME, is now a widely accepted method of VOC identification with published SOPs relating to the topic[307]. There have also been studies that have identified the canine ability to identify disease from the associated scent[308]. Analysis of human VOCs can be performed on exhaled breath, urine, sweat, skin, vaginal secretions, and faeces. The number of VOCs varies significant depending upon the source, with breath producing the most and blood the lowest[309].

There is strong anecdotal evidence that faeces from patients with GI disease have an abnormal smell, moreover, that different malodours are indicative of particular conditions. VOCs are the product of the microbiota and thus the specific changes that occur within it, in the presence of GI and other disease can be associated with changes in VOC profiles. Studies have demonstrated significant shifts in the gut microbiota in patients with colorectal cancer when compared to healthy controls, e.g. overrepresentation of the genera *Coriobacteriaceae, Roseburia, Fusobacterium* and *Faecalibacterium*. Interestingly, the genera *Roseburia*, *Fusobacterium* and *Faecalibacterium*, which are moderately enriched in tumours, belong to the major butyrate producing intestinal bacteria. Further supporting the theory that the protective nature of butyrate is only pertinent in the early phases of carcinogenesis[253].

Using faeces, breath and urine a number of studies have examined VOCs as potential biomarkers for colorectal cancer and pre-malignant lesions. They also used a number of different detection techniques including GCMS, canine sensing, nanosensors and “eNoses”. Peng *et al* used nanosensors linked to GCMS and examined breath in a number of cancer patients, including lung, breast and colorectal cancer. The exhaled breath of patients with colorectal cancer contained significantly higher levels of a number of VOCs, these were, 1, 10-(1-butenylidene) bis benzene, 1, 3-dimethyl benzene, 1-iodononane, 1, 1-dimethylethylthio acetic acid, 4-(4-propylcyclohexyl)-40-cyano [1, 10-biphenyl]-4-yl ester benzoic acid and 2-amino-5-isopropyl-8-methyl-1-azulene carbonitrile[310]. Altomare *et al* compared a pattern of 15 compounds in the breath of colorectal cancer patients and healthy people with a sensitivity of 86%, a specificity of 83 % and an accuracy of 85 % (area under the receiver operating characteristic (ROC) curve 0·852)[311].

Urine has been examined by Arasaradnam *et al* and Silva *et al*. Arasaradnam *et al* used Field Asymmetric Ion Mobility Spectrometer (FAIMS) to identify the VOCs. FAIMS analyses demonstrated that the VOC profiles of colorectal cancer patients were tightly clustered and could be distinguished from healthy controls. Sensitivity and specificity for colorectal cancer detection with FAIMS were 88% and 60% respectively[312]. Silva *et al* used SPME and GCMS to report 5 significantly elevated VOCs in the urine of patients with colorectal cancer. These were 4-methyl-2-heptane, hexanal, 3-heptanone and 1, 4, 5- trimethyl-naphthalene. The validity of which was questioned as they were thought to be environmental in origin, from petroleum products or lab contaminants[313,314]. Sonoda *et al* demonstrated the utility of VOC assessment by canine sensing. The dogs were able to identify the faeces from patients with colorectal cancer with a sensitivity of 97% and a specificity of 99%[308].

A study using a pattern recognition–based detection technique (eNose) looked at faecal VOCs for the identification of colorectal cancer and pre-cancerous adenomatous lesions. Faecal VOC profiles of patients with colorectal cancer differed significantly from controls. Receiver operating characteristic area under curve, 0.92; sensitivity, 0.85; and specificity, 0.87. Patients with advanced adenomas could also be distinguished from controls, with a receiver operating characteristic area under the curve of 0.79; sensitivity, 0.62; and specificity, 0.86[315].

The evidence suggests that VOC analysis for the detection of both colorectal cancer and pre-malignant colonic polyps could be utilised as a future screening tool. In isolation the evidence would suggest VOC superiority when compared to current methods, such as gFOBt. If used in combination with other methods of screening e.g. FIT, the accuracy of colorectal cancer-adenoma screening may be greatly improved. Improved sensitivities and specificities will lead to a higher yield of the screening programm, improved cost effectiveness and a reduction in unnecessary colonoscopies. If a suitable VOC based screening test could be identified using urine or breath then patient acceptability would be further improved in the absence of handling faeces.

Finally, when sensor technology progresses over the years, desktop or even consumer “eNoses” may become available to analyze VOCs, allowing for general practice based assessment or the study of changes in individual VOCs at home. This may be of specific importance when VOCs will be used in monitoring therapeutic response in treatment of colorectal cancer or to determine colonoscopic surveillance intervals[316].

**1.10 Hypothesis, aims and objectives**

The studies included will investigate the hypothesis that volatile organic compounds emitted from stool and urine differ in those with gastrointestinal malignancy, when compared to those without malignancy; allowing for the identification of VOC based biomarkers, paying particular attention to colorectal cancer and hepatocellular carcinoma. This difference will be categorised using HS-SPME-GC-MS.

The aim is to identify biomarkers that can be used for the screening and/or surveillance of colorectal cancer and hepatocellular carcinoma. It will also explore the utility of other biomarkers, namely faecal tM2-PK for the diagnosis of colorectal neoplasia.

The objectives are:

1) to further develop current methods for the analysis of urine using GCMS

2) to analyse urine samples from patients with hepatocellular carcinoma in order to identify potential biomarkers for screening and/or surveillance

3) to analyse urine samples from patients with colonic neoplasia in order to identify potential biomarkers for screening and/or surveillance

4) to analyse faecal samples from patients with colonic neoplasia in order to identify potential biomarkers for screening and/or surveillance

5) to determine whether the faecal tM-PK assay has a role in diverse settings or as an adjunct to existing FOBt based screening

6) to explore the potential carcinogenic role of the microbiome in colorectal cancer by examining the VOC profile of cultures *Fusobacterium nucleatum* and *Campylobacter showae*

# Chapter 2

# Standard operating procedure optimisation, sample acquisition, methods and statistical analysis

This chapter describes the development of standard operating procedures for the analysis of VOCs emitted from the headspace of human urine, paying particular attention to urine volume optimisation and preparation with freeze drying. In addition, laboratory methods for faecal and urine headspace SPME GCMS analysis, tM2-PK ELISA testing, headspace SPME GCMS analysis of *Fusobacterium nucleatum* and *Campylobacter showae* cultures are described. Finally, compound identification methods and statistical methods used in this thesis are also detailed.

## 2.1 Urine volume optimisation:

There is no generally accepted, standardised approach to optimise VOC yield from human urine. Many potential strategies, including using unadulterated urine, adding NaOH, HCl and freeze drying have been employed[317,318]. Within our laboratory, previous work had taken place looking at optimisation using 4ml of urine, in 10ml headspace vials, analysed by GCMS. It found that freeze drying was the method that yielded the most VOCs, whilst preserving the integrity of the GC column, together with a diverse range of VOCs[317]. The addition of acids or bases often results in an increase VOC yield but in favour of their respective chemical derivatives, however freeze srying does not appear to favour one chemical group giving a wider spread of compound class.

For the urine volume optimisation work, the HS-SPME-GCMS method used is fully described in section 2.3.

The urine samples provided by collaborators was supplied as volumes between 0.5ml and 1.5ml per patent, therefore experimental work was conducted in order to assess on optimal sample volume based upon the smaller volumes provided.

### 2.1.1 Experiment 1

Less than 4ml of urine was available for the present study. As a result, further work to optimise freeze drying of smaller urine volumes was performed. Freeze drying removes water from the urine samples and, therefore, concentrates the residual compounds.

### 2.1.2 Aim

To identify the optimal volume of urine for VOC analysis when using a freeze drying technique and 2ml headspace vials for GCMS.

### 2.1.3 Hypothesis

There will be a significant difference in the VOC yield from the different volumes of urine.

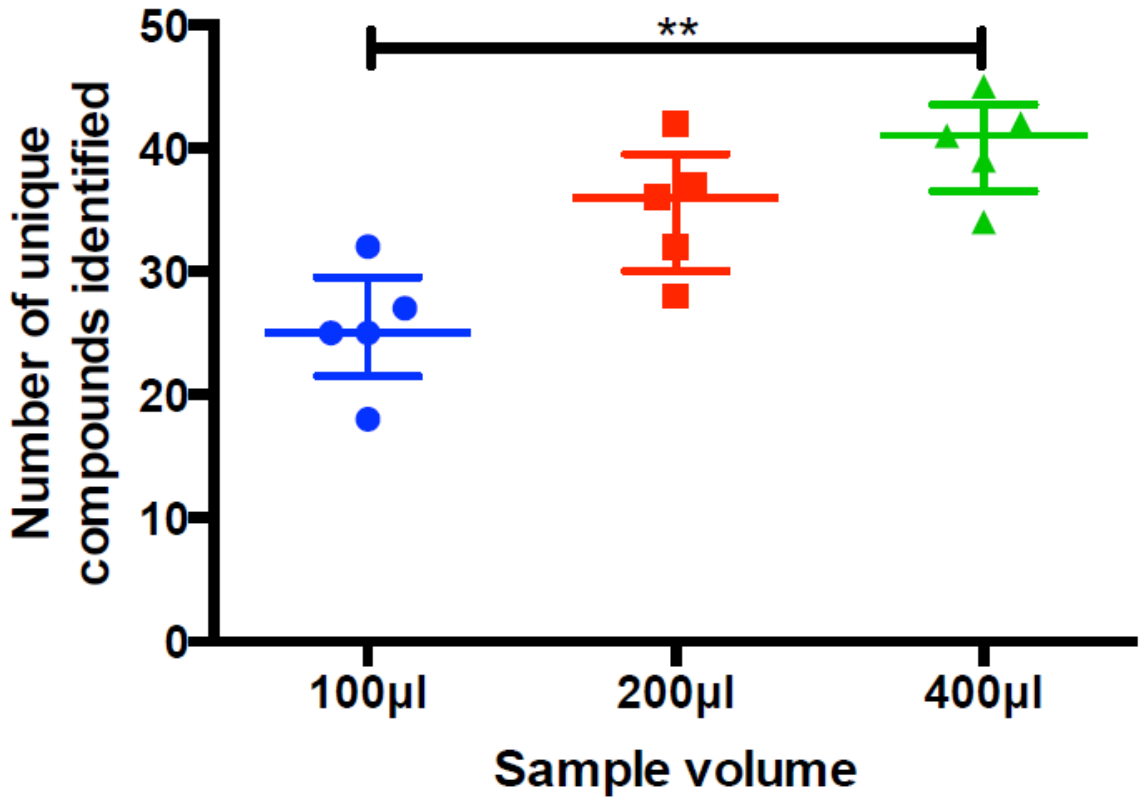
### 2.1.4 Experiment 1-Method:

First pass urine was collected from a healthy volunteer. The sample was collected in a 500 ml universal glass bottle from within the laboratory. This was then aliquoted as 100ul, 200ul and 400ul volumes into 2ml headspace vials (Agilent Technologies). Five replicates were made of each volume. These were then stored frozen at -20°C for 24 hours. Freeze drying was then conducted for a further 24 hours (-30°C and 8 mbar pressure), followed by storage at -20°C until analysis. GCMS, compound identification and statistical analysis methods described in section 2.2.1 were used to complete analysis.

#### 2.1.4.1 Results

100ul volumes of urine gave the fewest VOCs, mean= 25.4 (range 18-32). The 400ul samples gave the most VOCs, mean = 40.2 (range 34-45). Falling between these two were the 200ul samples, mean=35 (range 28-42).

**Figure 8: Scatter plot showing number of compounds identified for each sample and volume of urine.**




The mean and standard error of the mean bars are shown. ** P=<0.01 Kruskall-Wallis 1-way ANOVA.

#### 2.1.4.2 Conclusion

400ul appears to give the highest prevalence of VOCs, particularly when compared to the 100ul samples. There is no statistically significant difference between 200 and 400ul, despite this the mean VOC prevalence in the 400ul samples is larger than the 200ul. Based on this information, a further experiment was performed to compare the VOC yield from 400ul of urine freeze dried in a 2ml headspace vial to 4ml of urine freeze dried in a 10ml headspace vial.

### 2.1.5 Experiment 2

Experiment 1 had identified 400ul as the optimal volume of urine to use in a 2ml headspace vial when using the freeze dry technique. Previous work in our laboratory looking at urinary VOCs related to prostate cancer had used 750ul of urine in a 10ml headspace vial with NaOH added to the urine before GCMS analysis[318]. Later work had demonstrated the superiority of freeze drying and the use of 4ml of urine[317], with a high number of VOCs whilst maintaining a wide diversity of chemical class. In future larger volumes of urine will be available to allow the use of 4ml in a 10ml headspace vial. Identification of the difference, if any, between the 400ul and 4ml volumes was an important outcome of this experiment.

#### 2.1.5.1 Aim

To compare the number of VOCs identified from 400ul of freeze dried urine in 2ml headspace vial and 4ml of freeze dried urine in a 10ml headspace vial.

#### 2.1.5.2 Hypothesis

4ml of urine in a 10ml headspace vial will provide a greater number of VOCs than the 400ul in the 2ml headspace vials.

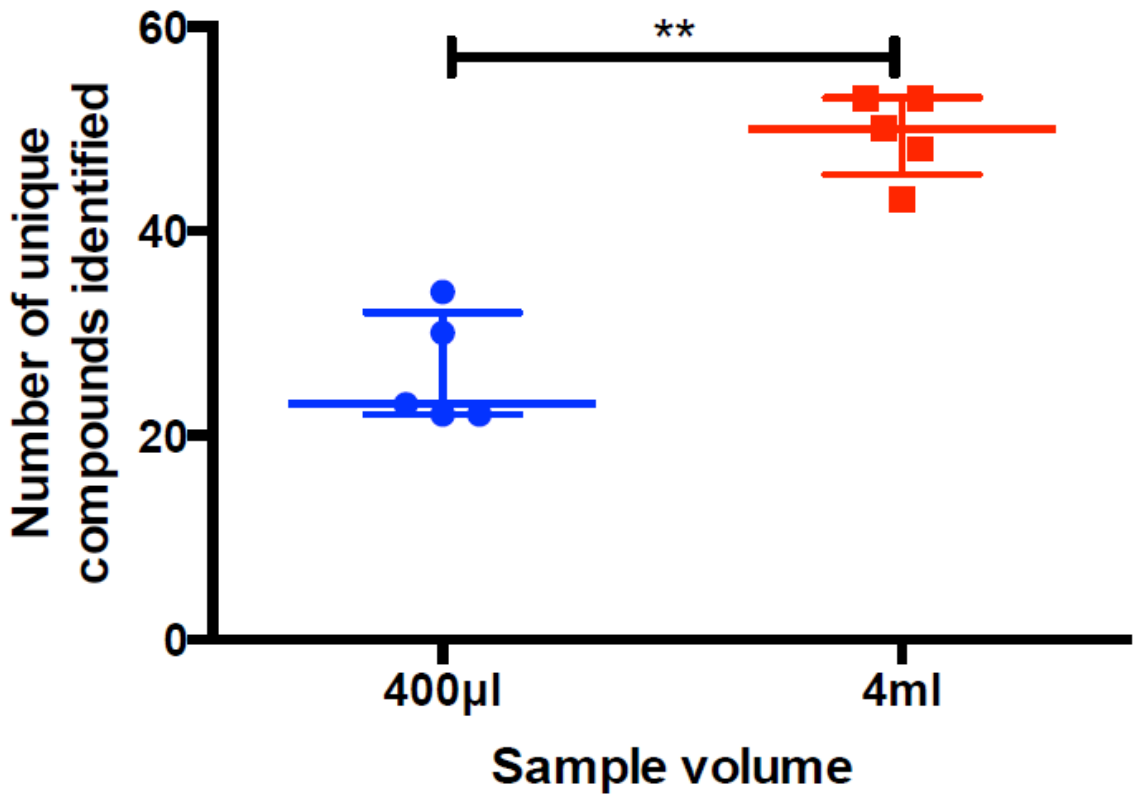
#### 2.1.5.3 Method

First pass urine was collected from the same volunteer. 400ul was then aliquoted in to five 2ml headspace vials (Agilent Technologies) and frozen at -20°C. Similarly 4ml was aliquoted into 10ml headspace vials. There were 5 replicates of each vial volume. After freezing at -80°C they were placed into the freeze dryer at -30°C and 8 mbar of pressure for 24 hours. Following this the samples were placed at -20°C for storage. GCMS, compound identification and statistical analysis methods are described elsewhere (sections 2.3 and 2.3.2).

#### 2.1.5.4 Results

There were significantly more VOCs identified in the 4ml samples, p=0.01.

**Figure 9: Scatter plot demonstrating the number of compounds identified in 400ul and 4ml urine samples.**




The standard error of mean bar and mean bar are shown, p=0.01.

**2.1.5.5 Conclusion:**

4ml of urine in 10ml headspace vial gives a better yield than smaller volumes in the 2ml headspace vial. However, if smaller volumes of urine are available analysis of 400ul in a 2ml headspace vials, that has been freeze dried, can be an option that yields a suitable number of VOCs. There was a difference in the mean number of compounds identified from the 400ul samples between experiment 1 and 2. Experiment 1 had a mean of 40 VOCs, whilst experiment 2 produced a mean of 24. The same GCMS method was used in both experiments, together with the same 2ml headspace vial. The urine utilised was different between the two experiments, it was from the same donor but acquired from two separate donation 1 week apart. The difference in the means is likely to originate from the variation in diet, microbiome and metabolic activity of the two donations.

## 2.2 Sample acquisition and handling

### 2.2.1 Urine acquisition and sample preparation for VOC analysis in hepatocellular carcinoma study

Acquisition of urine occurred in different locations for the small and larger pilot studies. Those forming the small pilot study were supplied by collaborators having been provided by a third party. This was permitted in line with existing ethical approval. Patient recruitment and sample acquisition for patient in Liverpool was in line with full ethical permission, ref 15/LO/0836.

As part of the small pilot study urine was collected from patients attending the hepatology service, at Queen Elizabeth Hospital Birmingham; they were initially collected in standard universal (hard plastic) containers. As soon as possible (in Birmingham) up to 1.5ml of the urine was aliquoted into an Eppendorf tube and stored at -20°C until use. All patients had a confirmed diagnosis of liver cirrhosis, based on a combination of clinical, biochemical and radiological markers; this included those with a diagnosis of hepatocellular carcinoma (HCC). Samples were then transferred, frozen, to Liverpool and stored at -20°C once again. Prior to preparation for analysis, there were no other freeze-thaw cycles. In Birmingham, the sex and age of the patient were recorded, along with the aetiology of the cirrhosis.

Before analysis, the samples were re-aliquoted, in batches of 10 to reduce the time out of the freezer; 400ul of urine was added to 2ml headspace vials (Agilent Technologies) and then refrozen before the next batch of 10 were removed from the freezer to be aliquoted. Once all 60 samples had been aliquoted they were stored at -20°C for a further 24 hours. All 60 samples were then subjected to freeze drying as a complete batch for 24 hours. Freeze drying was performed at -30°C and 8 mbar of pressure. Freeze drying was completed and the vials were sealed with magnetic caps of 2 ml (Agilent Technologies) and then stored at -20°C until processing[317].

Urine collection for the pilot study occurred at two clinical sites within the city of Liverpool. A random pass urine sample was provided by patients attending outpatient clinics at both sites and collected in standard universal containers. Clinical data was recorded and correlated. These samples were then returned to the laboratory and frozen at -20°C. Subsequently they were thawed and aliquoted as 4ml into a 10ml glass headspace vial and sealed with magnetic caps (Supelco, UK). Freeze drying was performed at -30°C and 8 mbar of pressure for 24 hours, the samples then being returned to -20°C storage until further steps were performed[317].

### 2.2.2 Stool acquisition and sample preparation for VOC analysis in colorectal cancer study

Patients undergoing colonoscopy, following a positive FOBt, as part of the Merseyside and Wirral Bowel Cancer Screening Programme (BCSP) were recruited. Additional symptomatic patients were

recruited from those undergoing colonoscopy for polyp surveillance, planned polypectomy, colonoscopy for the investigation of iron deficiency anaemia (IDA), change in bowel habit or abnormal radiological imaging; the FOBt of these controls was unknown. These patients were identified on the basis of the referral sent to the Endoscopy Department at the Royal Liverpool University Hospital. Patients with a confirmed diagnosis of colorectal cancer awaiting surgical resection or other management, at the Royal Liverpool University Hospital, were also included in the recruitment. These patients had not received any prior therapy for the cancer, including chemotherapy or radiotherapy. If these patients were identified via the BCSP then the FOBt status was noted as positive. Sample acquisition took place with full written consent, with BCSP committee and research ethics committee approval (14/SW/1162). Ethically approval was provided as amendments to existing local ethical approval in both Sheffield and Plymouth for the use of their samples.

BCSP patients were recruited when they attended their routine pre-procedural interview. Patient information sheets, consent forms and collection kits were distributed during this interview. Those undergoing colonoscopy for other indications were sent these items in the post with the other procedural items. Patients with a known diagnosis of colorectal cancer were recruited at their clinic appointments, during which time they were given the patient information sheet, consent form and collection kit.

Two other stool acquisition streams provided samples. The first was derived from all-comers to colonoscopy in Sheffield. Samples were provided by the patients and the results of their colonoscopy were retrospectively collected. The second stream, from patients in Plymouth, was obtained from patients undergoing colonoscopy investigating iron deficiency anaemia (IDA). Again the outcome of the respective colonoscopy was retrospectively recorded. The FOBt status of these patients was unknown.

During the recruitment, screening suitability and demographic questions were asked. This included smoking status and exposure to antibiotics in the last 6 months. Sample collection kits were provided to each patient (OdoReader Sampler, University of the West of England). This contained a hard plastic spoon to aliquot the sample into a dedicated glass vial. Patients were instructed to place at least 3 spoonfuls of faeces into the glass vial before sealing. Patients produced the sample within 48 hours of attending their colonoscopy and before commencing the required bowel preparation. They were asked to keep the sample in a cold environment prior to their attendance. The sample was then collected from the Endoscopy Department when the patients attended for their procedure.

Following transfer of the sample to the laboratory, 450mg of unadulterated faeces was then aliquoted into 10ml headspace vials and sealed with magnetic caps (Supelco, UK)[307]. Both the sample intended for analysis and the residual faeces were then stored at -20°C until GCMS analysis was performed.

### 2.2.3 Urine acquisition and sample preparation for VOC analysis in colorectal cancer study

Patients providing stool samples were also asked to provide urine samples. Therefore, the method of identification and patient recruitment was the same as that used in Section 2.2.2. A 10ml hard plastic universal container was provided to patients for collection of the urine, they were asked to fill the container with their sample. Patients produced the sample within 48 hours of attending their colonoscopy and before commencing the required bowel preparation. They were asked to keep the sample in a cold environment prior to their attendance. The sample was then collected from the endoscopy department upon the patient's presentation.

Following transfer to the lab 4ml of unadulterated urine was then aliquoted into 10ml headspace vials (Supelco, UK). This was subsequently frozen at -20° C, in batches of 20, because of the capacity of the equipment, the samples were then freeze dried at -30°C and 8 bar pressure for 24 hours. The vials were then sealed with magnetic caps and placed at -20°C until GCMS processing[317].

### 2.2.4 *Fusobacterium nucleatum* and *Campylobacter showae* culture

This was performed by Dr Georgina Hold, a collaborator in Boston, USA. It was performed as part of a larger piece of work between Harvard University and University of Aberdeen. Prior to the VOC analysis the *Fusobacterium nucleatum* and *Campylobacter showae* were maintained on Columbia blood plates (Columbia base agar + 10% sheep blood and with further supplementation of sodium formate (0.2% w/v) and sodium fumarate (0.3% w/v), under strict anaerobic conditions at 37°C. For the VOC experiments the bacteria were cultured in Columbia base broth + 10% foetal bovine serum and with further supplementation of sodium formate (0.2% w/v) and sodium fumarate (0.3% w/v), under strict anaerobic conditions at 37°C. To generate liquid cultures bacteria were transferred to liquid media and grown over night in mono-cultures. This occurred in 15 ml culture tubes. In the morning, cultures were measured to calculate OD (600nm) and diluted to an OD of 0.15 to generate equivalent cfu for both *Fusobacterium nucleatum* and *Campylobacter showae*. Previous work had confirmed using plate counts that an equivalent cfu was obtained from both organisms with an OD of 0.15 being equivalent to a cfu of $10^7$cfu/ml. Further confirmation was obtained by using qPCR analysis of 16S rRNA gene copies equivalents using universal and genus specific primers to allow comparisons of *Fusobacterium nucleatum* and *Campylobacter showae* levels within mixed cultures.

### 2.2.5 Faecal tM2-PK quantification

Stool samples provided by patients according to the method described in Section 2.2.2 were used, storage of the stool samples was as described in Section 2.2.2. ScheBo® Tumor M2PK™ stool ELISA kits, in conjunction with ScheBo® Master Quick-Prep™ for extraction (Biotech, Germany) were used for stool analysis. An M2PK cut-off of 4u/ml was used. Four training events/sessions were made using ready-to-use standards ensuring operator technique prior to analysis of samples. No more than 18 samples were analysed on a single ELISA plate in order to minimise incubation time between the first and last sample. Duplicates were carried out for all samples (Figure 10).

**Figure 10: ELISA plate layout for M2PK analysis.**

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| A | Blank | Blank | Sample XX | Sample XX | Sample XX | Sample XX |
| B | Standard 1 | Standard 1 | Sample XX | Sample XX | Sample XX | Sample XX |
| C | Standard 2 | Standard 2 | Sample XX | Sample XX | Sample XX | Sample XX |
| D | Standard 3 | Standard 3 | Sample XX | Sample XX | Sample XX | Sample XX |
| E | Standard 4 | Standard 4 | Sample XX | Sample XX | Sample XX | Sample XX |
| F | Control | Control | Sample XX | Sample XX | Sample XX | Sample XX |
| G | Sample XX | Sample XX | Sample XX | Sample XX | Sample XX | Sample XX |
| H | Sample XX | Sample XX | Sample XX | Sample XX | Sample XX | Sample XX |

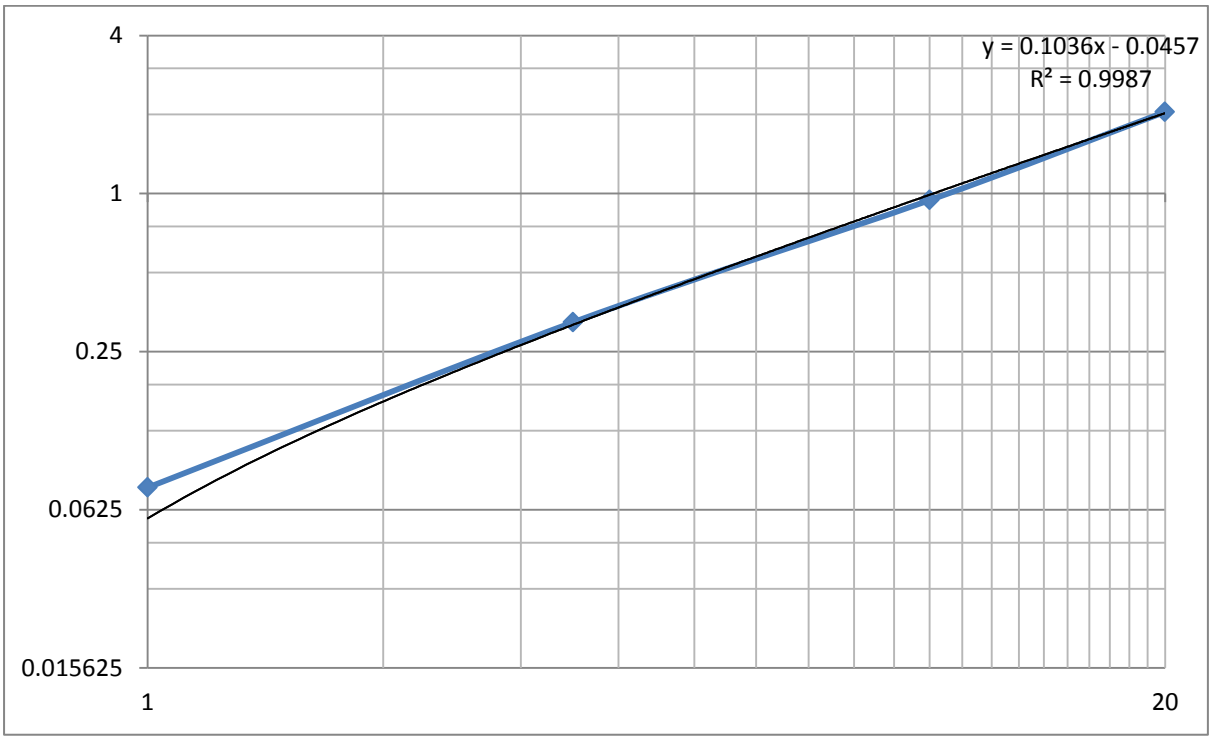
Optimisation of operator technique used columns 1 and 2, together with row A-F. Plates 1-5 contained 18 samples, the sixth plate contained 10 samples.

A maximum of 18 samples were randomly removed from frozen storage and left to thaw at room temperature for one hour. The ScheBo® Master Quick-Prep™ extraction kits, with their dosing tip, allowed for a reproducible extraction process, with a previously documented coefficient of variance of 2.7%. Each extraction tube was vortexing and left to stand allowing the sediment to settle. This

process took an average of 20 minutes. During this time the investigator was blind to the diagnosis of each sample. The diagnosis was only correlated to the samples when the analysis was complete.

A ready to use substrate solution was used for the colour reaction followed by the included stop solution. Optical density was read at 450nm using an automated plate reader, 5 minutes after the stop solution had been applied. A standard curve was generated for each plate using the optical densities of the internal standards, plotted on log-log axes. This analytical performance of each plate was assessed by interpolated control calculation using a sample of known tM2-PK abundance. If a deviation >15 % from expected concentration was identified assay failed QC and was repeated.

**Figure 11: Standard curve generated from one M2PK ELISA plate.**




**2.3 Headspace extraction and gas chromatography mass spectrometry method**

The same method was applied to urine and stool and includes samples analysed for colorectal cancer and hepatocellular carcinoma. As described all samples were placed in 10ml headspace vials (Supelco, UK).

A Perkin Elmer Clarus 500 GC/MS quadruple bench top system (Beaconsfield, UK) was used in combination with a Combi PAL autosampler (CTC Analytics, Switzerland) for the analysis of all samples. SPME-GC-MS method was used. The GC column used was a Zebron ZB-624 with inner diameter 0.25 mm, length 60 m, film thickness 1.4 µm (Phenomenex, Macclesfield, UK). The carrier

gas used was helium of 99.996% purity (BOC, Sheffield, UK). The SPME fibres used were CAR-PDMS 85 µm (Sigma-Aldrich, Dorset, UK).

Samples were placed on the autosampler tray in batches of five, each batch being preceded and followed by lab air assessment to monitor for contaminates. During all processing the tray was kept at a constant 1°C. This was performed as a continuous process in order to maximise efficiency. Once 10 samples were available following recruitment they were processed through the GCMS in order to minimise time frozen at -20°C. Specifically, this minimised the time difference between freezing the first sample obtained and processing the last.

The fibre desorption conditions were 5 minutes at 220°C. The initial temperature of the GC oven was set at 40°C and held for 1 minute before increasing to 220°C at a rate of 5°C/min and held for 4 min with a total run time of 41 min. A solvent delay was set for the first 6 min and the MS was operated in electron impact ionization EI+ mode, scanning from ion mass fragments 10 to 300 *m/z* with an interscan delay of 0.1 sec and a resolution of 1000 at FWHM (Full Width at Half Maximum). The helium gas flow rate was set at 1 ml/min. The sensitivity of the instrument was determined with 2-pentanone only and will vary for other compounds[307,318,319].

### 2.3.1 Headspace extraction and gas chromatography mass spectrometry method relating to *F. nucleatum* and *C. showae*

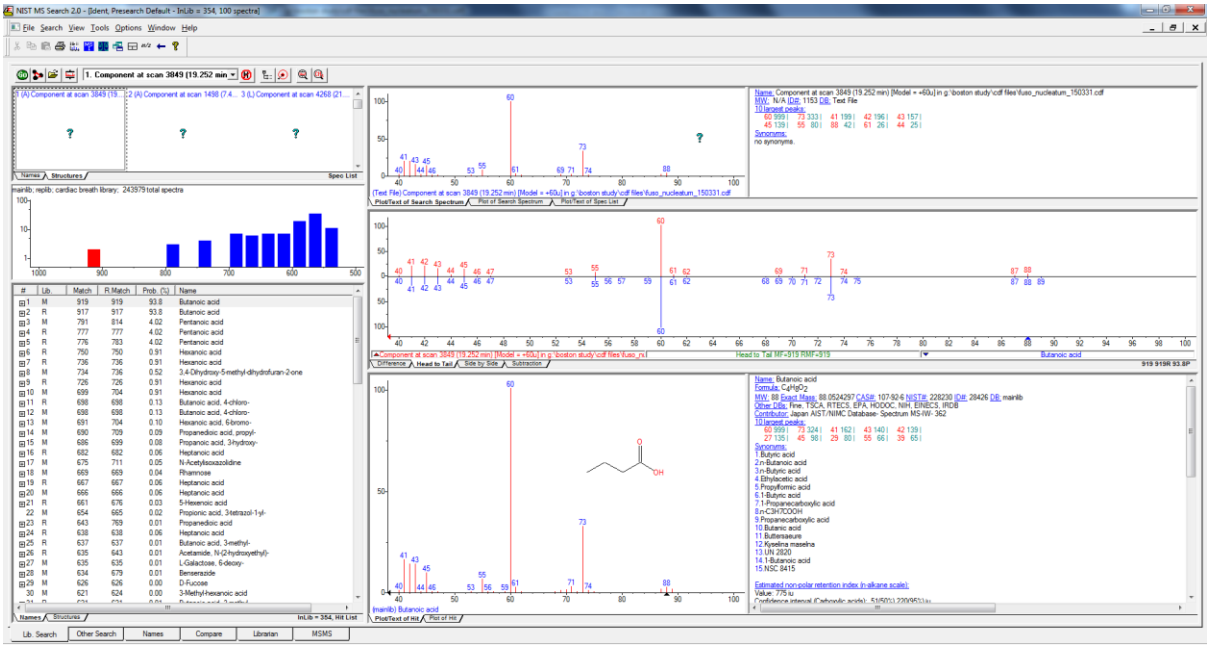
A manual injection technique was used for this element of the work. Specific manual SPME fibres were used DVB/CAR-PDMS (Sigma-Aldrich, Dorset, UK) which underwent a 10 minute exposure to the headspace gas of the culture vials in Boston, USA. The fibres were shipped to Liverpool within 48 hours at ambient temperature. They were transported with a septum covering the inlet of the fibre. They were then placed at -20°C until processing. The Perkin Elmer Clarus 500 GC/MS quadruple bench top system (Beaconsfield, UK) was used with injection via the manual port. There were 35 samples in total which were processed as a continuous batch, with every third fibre being followed by a blank fibre, that had been stored in Liverpool only. The GCMS settings, were as above, Section 2.3.

### 2.3.2 Identification of volatiles

Mass spectral data were analysed using Automated Mass Spectral Deconvolution and Identification System (AMDIS) (AMDIS-version 2.71, 2012) and the National Institute of Standards and Technology (NIST) (version 2.0, 2011). Data bridge was used to covert raw data files into CDF files that could then be analysed in AMDIS. Three components were used within NIST to identify the compounds,

applying percentage probability, a forward and reverse matching of 800/1000 as above, allowing for compound identifications (Figure 12). A compound was considered to be present when it satisfied these 3 criteria. This process provides a relative ion abundance, therefore no units of ion abundance are available. Further processing and alignment of data was performed using the Metab package in R[320]. A dedicated library for analysis of these samples was built using at least 50% of the samples in each cohort. During this library building the investigator was blind to the sample diagnosis or cultured bacterium. Log transformation was used to normalise data.

**Figure 12: Screenshot of the NIST VOC library used to identify compounds.**



The image includes the "match" process, probability assessment and head-to-tail fragment analysis. This is the screen used to identify all compounds and in this case butanoic acid.

### 2.4 Data handling and statistical analysis

#### 2.4.1 GCMS data handling and statistical analysis

All GC-MS data was processed using the AMDIS (version 2.71, 2012) in conjunction with the NIST mass spectral library (version 2.0, 2011) and the R package Metab[320]. VOCs were identified using an in-house library built from a combination of previous work on urine GCMS and at least 50% of the samples in each study. AMDIS in combination with the NIST library was utilised for this element. All statistics were performed using R version 3.0.2[90,321,322]. Fisher's exact test, Student's t test, chi-squared testing and a one-way analysis of variance (ANOVA) followed by HSD Tukey testing was

applied. Principal component analysis (PCA) was used to show similarities within data classes. P-values <0.05 were considered significant (unless specified otherwise). Finally, partial least squares discriminant analysis (PLS-DA) was used to look for separation within the samples followed by ROC analysis testing of potential biomarker candidates, using Metaboanalyst(3.0)[323,324]. Within Metaboanalyst relative abundances were normalised by log transformation and auto-scaling, an example of this can be been in the Appendix 4. This consists of a data output report of the samples used in the studies detailed within.

### 2.4.2 Biomarker analysis of VOCs

Analysis of potential biomarkers was conducted using Metaboanalyst[324]. Following the assessment of fold change and Student's t test, potential VOCs were taken forward for biomarker analysis. This was performed in Metaboanalyst and was performed on quantitative data. The VOC abundance was normalised by median, log transformed and scaled using the auto-scaling function built into Metaboanalyst. Initially univariate ROC analysis was performed. Combination testing was based upon previous p values and FDR values. Calculating ratios of all possible metabolite pairs and then choosing top ranked ratios, based on p values, allowed for further biomarker assessment. Logistic regression and a 10-fold cross validation modelling was used to generate AUROC, sensitivity and specificity[324]. An example of the Metaboanalyst data handling process can reviewed in Appendix 4.

### 2.4.3 tM2-PK data handling and statistical analysis

The optical reader was programmed to read three separate optical densities for each plate, this was performed at 8 second intervals. A mean optical density from these three runs could then be calculated. Standards provided with each ELISA kit were used to generate the standard curve and $R^2$ value for each plate. These were correlated with the 4.5u/ml control, +/- 15%. Once optical densities had been produced each sample was then cross referenced to the colonoscopy findings cancer, adenoma or normal. Using a pre-defined cut off of 4u/ml, each M2PK result was defined as true positive (TP), false positive (FP), true negative (TN) and false negative (FN), according to the colonoscopy findings. All analysis was performed by an investigator blinded to the patients' diagnoses, under standardised conditions. Graphpad Prism 6 was used to compare pathological groups by two tailed t test and Kruskal-Wallis and Dunn's post-hoc analyses (p value significant <0.05), and to determine the area under ROC curves, sensitivity, specificity, positive predictive value and negative predictive values. Further testing was performed using a binomial method McNemar's test.

# Chapter 3

# Urinary volatile organic compounds as a biomarker for the diagnosis of hepatocellular carcinoma

### 3.1 Introduction

Primary liver cancer/hepatocellular carcinoma ( HCC) is the third largest contributor to cancer mortality in the world[122]. Mortality and 5-year survival from HCC is variable worldwide, but is considered to be poor. HCC typically occurs on a background of cirrhosis, developing as small asymptomatic nodules. It has an average doubling time of 6 months. The size of the HCC and severity of liver disease at the point of diagnosis have a significant impact on survival, therefore early diagnosis and screening should be beneficial. Current European guidelines recommend 6-monthly liver ultrasound scanning of patients with cirrhosis to screen for HCC. Detection of early lesions by ultrasonography has been reported to be as low as 63%, therefore the addition of other screening modalities has the potential to improve this[144].

#### 3.1.1 Aim

The study was performed as two parts, a small pilot study followed by a larger scale replication study.

1) The aim of the small (Birmingham) pilot study was to assess a recently published departmental standard operating procedure to a real life cohort and then to assess if differences in VOCs emitted from urine could be detected between those with and without HCC[317].

2) The aim of the larger (Liverpool) replication study was to assess the utility of these differences in urinary based VOCs as a biomarker for the diagnosis and/or monitoring of HCC.

#### 3.1.2 Method

For the small pilot study, samples were obtained from Prof Philip Johnson, who had recruited patients from University Hospitals Birmingham NHS Trust. SPME headspace extraction followed by GCMS was used to assess VOCs emitted from freeze dried urine of cirrhotic patients with and without HCC. Four hundred microlitres of urine was freeze dried in 2ml headspace vials for the small pilot study. After further method development, the larger pilot used 4ml of freeze dried urine to perform SPME GCMS in the larger pilot these were supplied by patients from Liverpool, UK: the full method is described in Chapter 2. Data analysis was performed in R[320] and Metaboanalyst[324], utilising Student's t test, Fisher's exact test, false discovery rate correction, PLS-DA and ROC analysis. Logistic regression modelling with 10-fold cross validation was used to test potential biomarkers.

### 3.2 Results

#### 3.2.1 Small pilot study

In total 62 samples were included in this study collected from 29 patients with HCC and 33 without HCC. Patients were HCC were older than those without (p= 0.02). There was no significant difference with regard to gender (p= 0.1) (Table 5).

**Table 5: Demographic details of patients included in the small pilot study.**

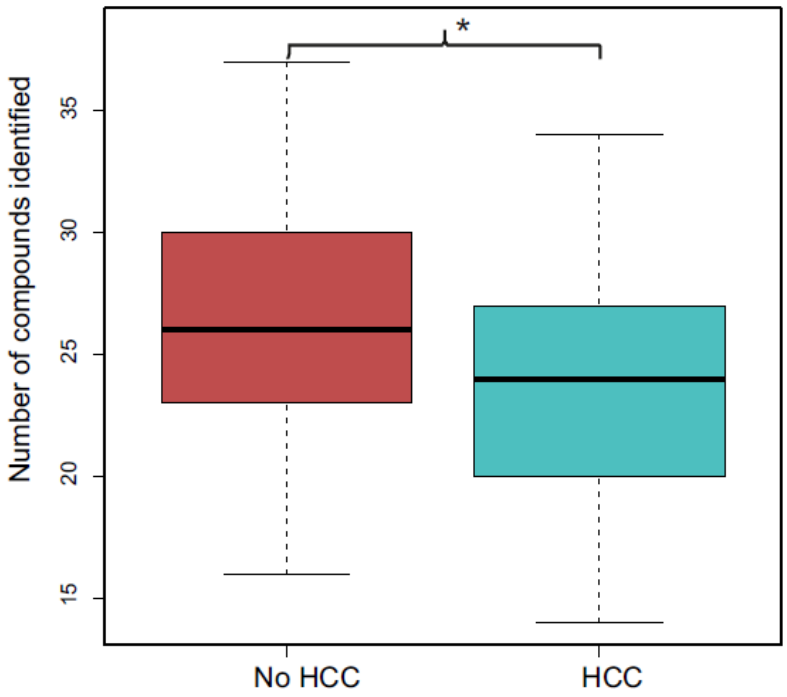
| | Male (%) | Female (%) | Mean age, years (range) |
|---|---|---|---|
| **No HCC** | 19 (58%) | 14 (42%) | 61 (35-82) |
| **HCC** | 20 (74%) | 7 (26%) | 68 (46-78) |

A total of 118 VOCs were identified across all samples. The mean number of VOCs in all samples was 25.1. There were significantly more VOCs identified in the cirrhotics without HCC compared with those with HCC, p=0.034 (Table 6, Figure 13).

**Table 6: Mean number of VOCs identified in each group, with standard deviation and error of mean.**

| | Cirrhosis without HCC | Cirrhosis with HCC |
|---|---|---|
| **Mean number of VOCS** | 26.5 | 23.7 |
| **Standard deviation** | 5.2 | 4.9 |
| **Standard error of mean** | 0.9 | 0.9 |

**Figure 13: Box plot to show the distribution of the number of VOCs identified in the patients without and those with HCC**




*p=<0.05

Categorical comparison for the prevalence of VOCs was performed and found one VOC to be signficantly different: methanethiol, whichwas found in 19/33 (57.6%) cirrhotics without HCC and 6/27(22.2%) with HCC (p=0.0083). Semi-quantitative analysis demonstrated 4 VOCs, including methanethiol, the abundance of which was significantly greater in those without HCC (p=<0.05), Table 3. However, when false discovery rate correction was applied in order to correct for multiple comparisons, none of the 4 VOCs achieved a significant q value (≤0.05).

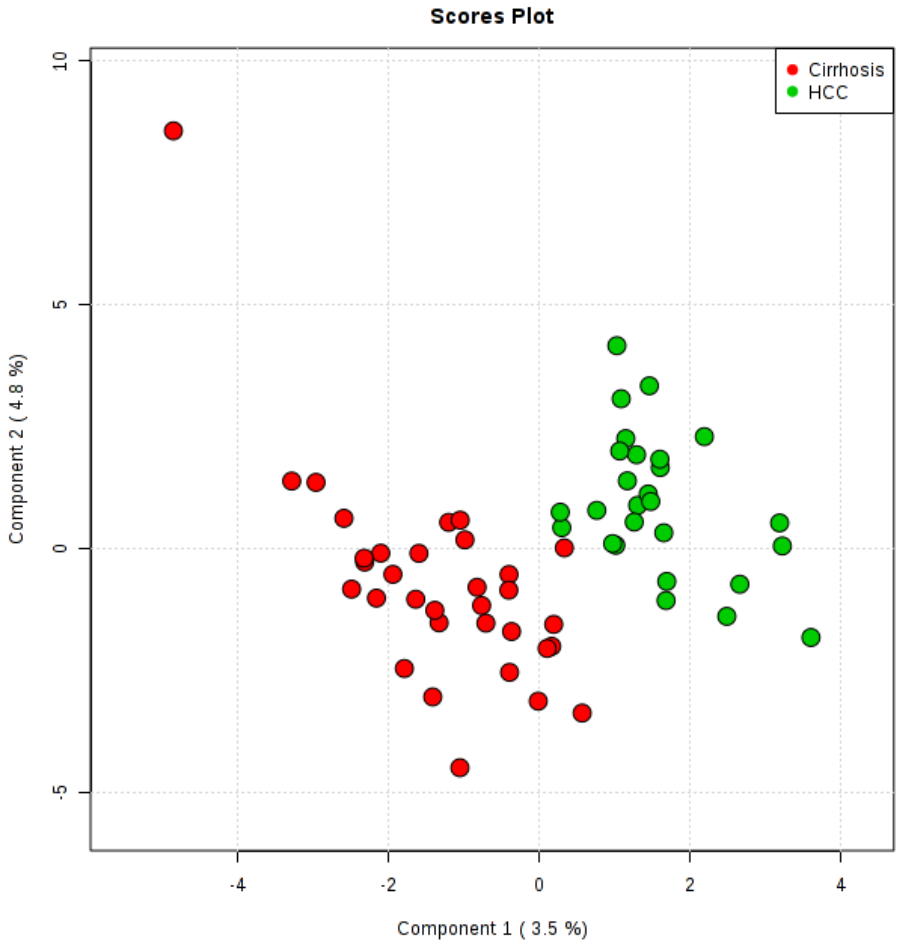
PLS-DA comparing those with and without HCC showed separation suggesting a potential discriminant utility for VOC analysis (Figure 14).

**Table 7: VOCs shown to be different following assessment of abundance using Metaboanalyst.**

| VOC | p value | False discovery rate q value |
|---|---|---|
| **Methanethiol** | 0.005 | 0.6 |
| **Dimethyltrisulphide** | 0.01 | 0.6 |
| **2,2-dimethylbutanoic acid** | 0.02 | 0.6 |
| **2,2,2-trifluoroethanol** | 0.02 | 0.6 |

False discovery rate applied in order to correct for multiple comparisons (q=<0.05 deemed significant).

**Figure 14: PLS-DA generated using all VOCs in all the samples from the smaller (Birmingham) pilot study.**

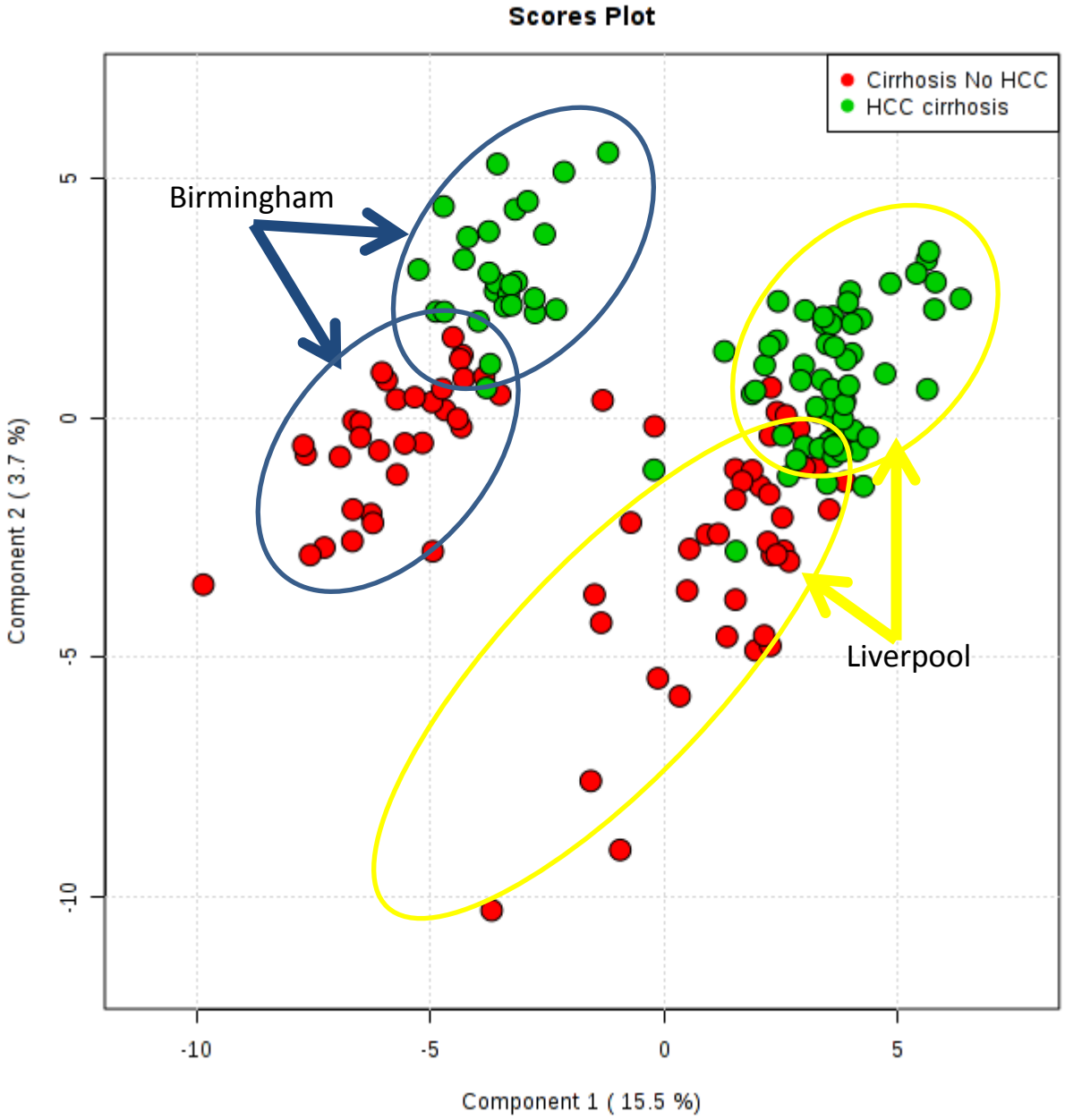


### 3.2.2 Diagnostic model and validation

The 4 VOCs identified in the smaller pilot study were then entered into a logistic regression model and 10 fold cross validation used to perform ROC analysis. The optimal AUROC achieved was 0.75 (95% CI 0.61-0.85), sensitivity 85% (95% CI 0.8-0.89) and specificity 58% (0.52-0.63). Validation of this model failed when applied to the Liverpool cohort from the pilot study. When the Birmingham VOCs were applied to the Liverpool samples the AUROC fell from 0.75 to 0.47.

PLS-DA examining all samples from both the small pilot and larger pilot study suggested a significant intra and inter cohort difference (Figure 15).

**Figure 15: PLS-DA containing all samples from both pilot studies.**




Clustering of samples highlighted by blue circles (Birmingham) and yellow circles (Liverpool).

### 3.3 Larger scale pilot Study

#### 3.3.1 Patient demographics

One hundred and three patients were recruited: 59 with HCC and 44 without. There was a male predominance in both the HCC and no HCC group. The mean age of the entire cohort was 64 years: patients with HCC were significantly older than those without, p=<0.0001, (Table 8). All patients had previously been diagnosed with cirrhosis as part of their clinical assessment and management. The diagnosis of cirrhosis was based upon clinical, biochemical and radiological parameters and was made by the managing clinicians, prior to recruitment.

**Table 8: Patient demographic details and clinical parameters for those included in the Liverpool pilot study analysis.**

| | No HCC | HCC | Total | p value |
|---|---|---|---|---|
| **Number** | 44 | 59 | 103 | - |
| **Mean age, years (range)** | 59 (40-78) | 68 (41-86) | 64 (40-86) | <0.0001 |
| **Gender** | | | | 0.05 |
| Male | 35 | 48 | 83 | |
| Female | 9 | 11 | 20 | |
| **Childs Pugh score** | | | | 0.6 |
| A | 33 | 44 | 77 | |
| B | 10 | 15 | 35 | |
| C | 1 | 0 | 1 | |
| **Average MELD score (range)** | 10.3 (7-23) | 9.6 (6-19) | 9.9 (6-23) | 0.2 |
| **Underlying aetiology** | | | | 0.4 |
| ALD | 16 | 27 | 43 | |
| Viral | 12 | 9 | 21 | |
| Metabolic disease | 14 | 20 | 34 | |
| Autoimmune | 2 | 3 | 5 | |

All patients included in the HCC group had had their diagnosis confirmed via the standard clinical methods and multi-disciplinary assessment. Of the 59 patients, 39 had received treatment: 38 in the

form of radiofrequency ablation (RFA) or trans-arterial chemoembolisation (TACE), these treated patients still had radiological evidence of active disease (Table 9).

**Table 9: Clinical parameters for those in the HCC cohort, stratified by treated and treatment naive.**

| | Treatment naive | Treated | Total | p value |
|---|---|---|---|---|
| **Number** | 20 | 39 | 59 | |
| **BCLC stage** | | | | 0.3 |
| A | 15 | 34 | 49 | |
| B | 4 | 3 | 7 | |
| C | 1 | 1 | 2 | |
| D | 0 | 1 | 1 | |
| **Mode of treatment** | | | | NA |
| TACE | NA | 26 | 26 | |
| RFA | NA | 10 | 10 | |
| TACE and RFA | NA | 2 | 2 | |
| Resection | NA | 1 | 1 | |
| Sorafenib | NA | 0 | 0 | |
| **Multifocal lesions** | | | | 0.02 |
| Yes | 6 | 24 | 30 | |
| No | 14 | 15 | 29 | |
| **Lesion > 3 cm** | 11 | 20 | 31 | 0.5 |
| **AFP positive** | | | | 0.7 |
| Yes | 8 | 14 | 22 | |
| No | 12 | 25 | 37 | |

*BCLC= Barcelona clinic liver cancer staging.*

*AFP cut off defined by trust based assay, as performed at RLUH and UHA.*

### 3.3.2 VOC profile of larger scale Liverpool cohort

Across all 103 samples, 216 VOCs were identified. There was no significant difference in the number of VOCs identified between the two groups, p=0.4 (Figure 16). The mean number of VOCs was greater in the larger pilot cohort (Table 10).

**Figure 16: Box plot to show the number of VOCs identified in the patients without and those with HCC from the larger pilot study cohort.**

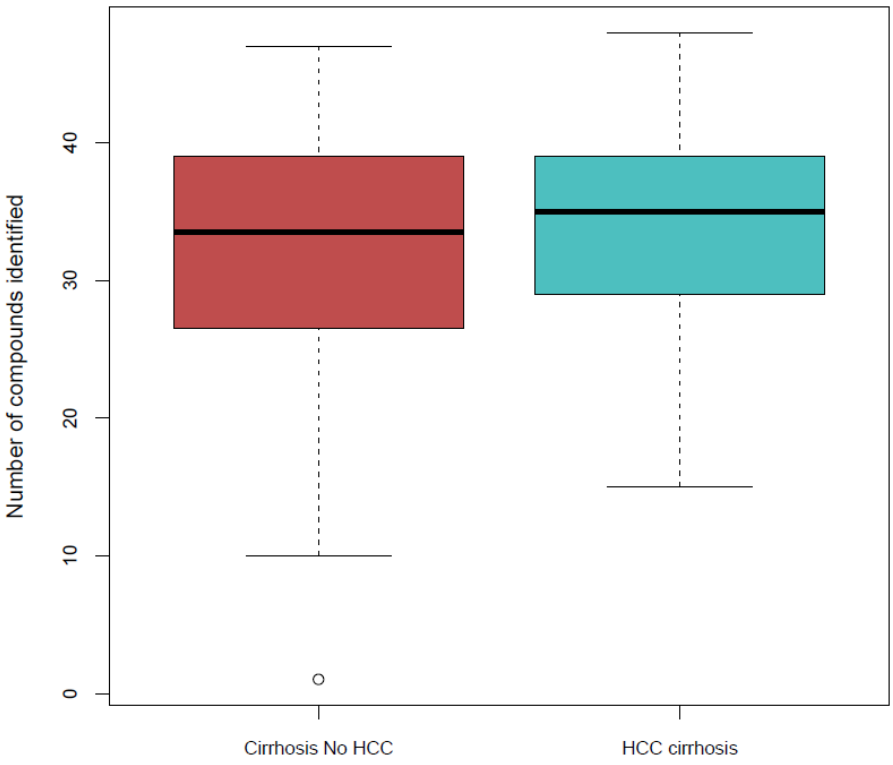



**Table 10: Mean number of VOCs identified in each group, with standard deviation and error of mean.**

| | No HCC | HCC |
|---|---|---|
| **Mean number** | 33.5 | 35 |
| **Standard deviation** | 9 | 7.6 |
| **Standard error of mean** | 1.3 | 0.9 |

### 3.4 Comparisons to assess and model biomarker utility

For all modelling and ROC analysis a logistic regression with 10-fold cross validation was used. Only significantly different VOCs with a p value of ≤0.01 were considered for candidate testing. All analysis in this section was conducted in Metaboanalyst and is quantitative[324].

#### 3.4.1 Comparison of VOCs in patients with cirrhosis and those complicated by HCC

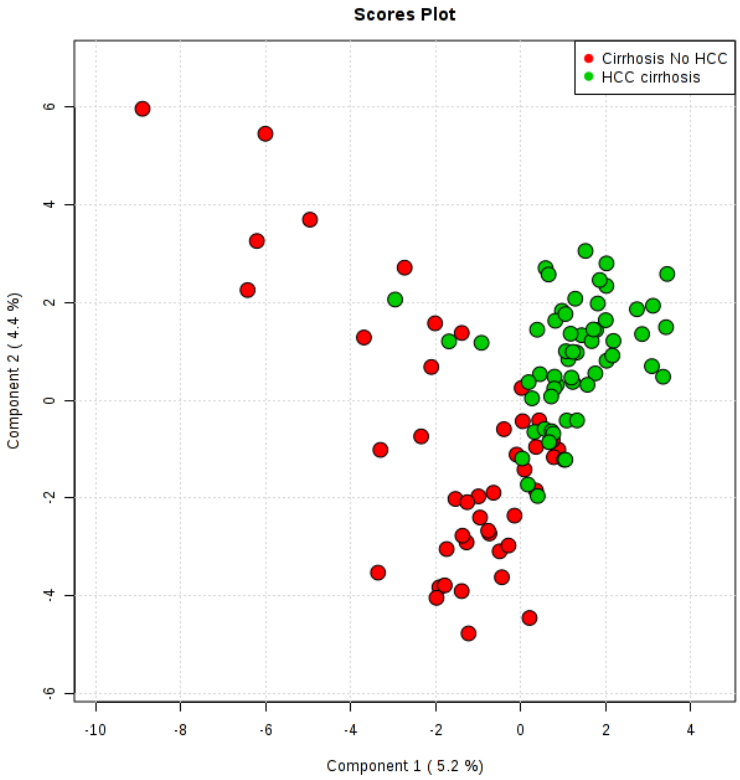
All 103 samples were assessed with 3 VOCs demonstrating a p value ≤0.01 (Table 9). PLS-DA based on all samples demonstrated separation, with particular clustering of the HCC cohort (Figure 6). ROC analysis based upon the 3 VOCs in Table 9 demonstrated an AUROC 0.76 (95% CI 0.65-0.83), sensitivity 65% (95% CI 0.61-0.69) and specificity 74% (95% CI 0.69-0.78).

**Table 11: VOCs found to be significantly different (p= ≤0.01), when assessing VOC abundance in those with and without HCC.**

| VOC | p value | False discovery rate |
|---|---|---|
| **Styrene** | <0.0001 | 0.03 |
| **Pyrazine-2-5-dimethyl-3** | 0.003 | 0.1 |
| **D-limonene** | 0.004 | 0.1 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

**Figure 17: PLS-DA based upon all VOC in all samples from the larger Liverpool pilot study comparing those with and without HCC.**




### 3.4.2 Comparison of VOCs of HCC patients who are treatment naive and who have received treatment

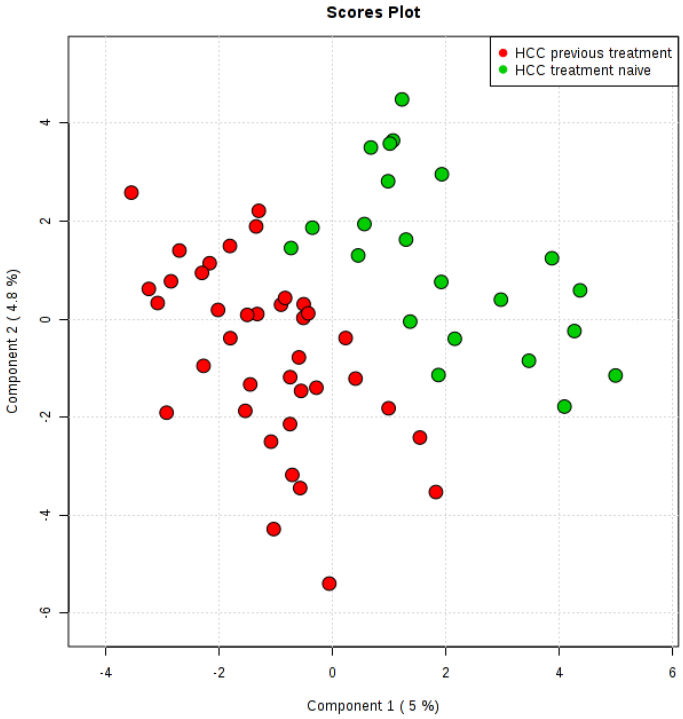
The 59 HCC patients were then stratified according to their treatment status. Two VOCs were found to be significantly different (Table 12). PLS-DA demonstrated separation of the cohorts with no overlap of samples (Figure 22). ROC analysis combining the two VOCs in Table 12 demonstrated an AUROC 0.75 (95% CI 0.51-0.89), sensitivity 73% (95% CI 0.67-0.79) and specificity 76% (95% CI 0.72-0.81).

**Table 12: VOCs found to be significantly different (p ≤0.01), when assessing VOC abundance in HCC patients who had received treatment and were treatment naive.**

| VOC | p value | False discovery rate |
|---|---|---|
| **Phenol** | <0.0001 | 0.04 |
| **3-methyl-2-butanone** | 0.001 | 0.1 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

**Figure 18: PLS-DA based upon all VOC in all samples when assessing VOC abundance in HCC patients who had received treatment and were treatment naive.**




### 3.4.3 Comparison of VOCs in cirrhotic patients without HCC and treatment naive HCC

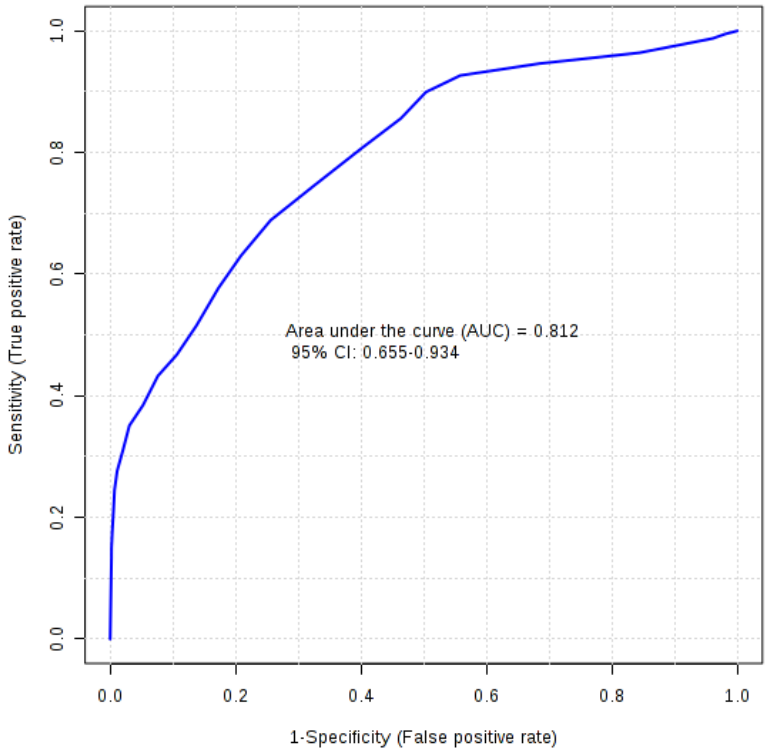
All 44 patients without HCC and the 20 treatment naive HCC patients were entered into this analysis. Five VOCs were found to be significantly different, with three - styrene, phenol and D-limonene being identified as significant in the earlier analysis (Table 13). Combining these VOCs as a biomarker panel for treatment naive HCC produced an AUROC of 0.81 (95% CI 0.66-0.93), sensitivity 77% (95% CI 0.71-0.83) and specificity 75% (95% CI 0.71-0.79) (Figure 19).

**Table 13: VOCs found to be significantly different (p= ≤0.01), when assessing VOC abundance in those without HCC and HCC patients who were treatment naive.**

| VOC | p value | False discovery rate |
|---|---|---|
| **Phenol** | 0.005 | 0.3 |
| **Styrene** | 0.006 | 0.3 |
| **D-limonene** | 0.01 | 0.3 |
| **Benzene-1-1-dimethylethoxy-** | 0.01 | 0.3 |
| **o-cymene** | 0.01 | 0.3 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

**Figure 19: ROC curve produced with logistic regression and 10-fold cross validation modelling, based upon the 5 VOC found to be significantly different for no HCC against treatment naive HCC.**




### 3.4.4 Comparisons of VOCs in HCC patients according to their Alpha-Fetaprotein status at diagnosis

Thirty percent of HCCs will not produce alpha feto-protein (AFP), therefore this 30% will be missed if AFP is used as a biomarker. This absence of production also suggests a difference in underlying metabolic pathways within the HCC.

#### 3.4.4.1 AFP negative versus AFP positive HCC

Thirty seven (62%) of the HCC patients in our cohort were AFP negative. When compared to those HCC patients who were AFP positive three VOCs were identified as being significantly different, p= ≤0.01, (Table 142). PLS-DA using all samples and all VOCs demonstrated separation.

**Table 14: VOCs found to be significantly different (p= ≤0.01), when assessing VOC abundance in those AFP negative and positive HCC.**

| VOC | p value | False discovery rate |
|---|---|---|
| **Ethylalcohol** | 0.001 | 0.1 |
| **3-methyl-2-hexanone** | 0.007 | 0.4 |
| **3-methylene-2-pentanone** | 0.01 | 0.4 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

### 3.4.4.2 Comparison of VOCs in those cirrhotics without HCC and those HCC patients with a negative AFP

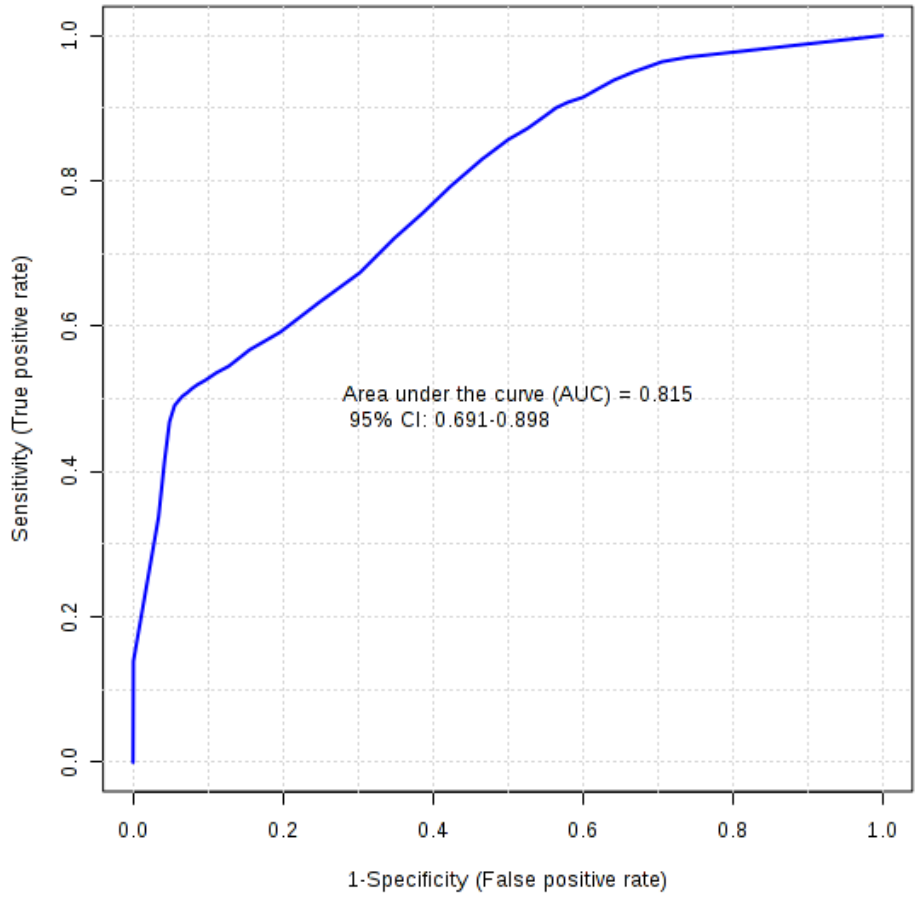
Comparison of all those without HCC and the 62% of our cohort with HCC and a negative AFP was then conducted, thus exploring the ability of VOCs to identify those HCCs that would not be identified by AFP. Five compounds were found to be significantly different, with styrene and D-limonene again being identified (Table 15). These compounds were entered into logistic regression and 10 fold cross validation ROC analysis in order to assess diagnostic utility, AUROC 0.82 (95% CI 0.69-0.89), sensitivity 84% (95% CI 0.8-0.88) and specificity 66% (95% CI 0.61-0.71) (Figure 20).

**Table 15: VOCs found to be significantly different (p= ≤0.01), when assessing VOC abundance in those with no HCC and AFP negative HCC.**

| VOC | p value | False discovery rate |
|---|---|---|
| **Styrene** | <0.0001 | 0.06 |
| **Ethylalcohol** | 0.001 | 0.06 |
| **D-limonene** | 0.001 | 0.07 |
| **2-3-butanedione** | 0.006 | 0.1 |
| **Pyrazine-2-5-dimethyl-3** | 0.01 | 0.3 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

**Figure 20: ROC curve produced with logistic regression and 10 fold cross validation modelling, based upon the 5 VOC found to be significantly different for no HCC against AFP negative HCC.**




### 3.4.5 VOC comparisons stratified by Barcelona Clinic Liver Cancer stage

The Barcelona Clinic Liver Cancer (BCLC) staging tool is commonly used as a decision aid in UK practice. The staging tool separates disease into early (potentially curative) disease and later (non-curative disease)[152].

Comparisons demonstrated significant differences in VOCs between those deemed to have curative (n=16) and non curative disease (n=43) according to the BCLC staging. Further sub-analysis was then performed to assess the ability of VOCs to identify those with early and therefore potentially curative disease, against a population of cirrhotics without HCC.

### 3.4.5.1 Comparison of VOCs in cirrhotic patients without HCC and those whose HCC is classified curative by BCLC

This comparison showed five VOCs to be significantly different, with styrene, D-limonene, 3-methyl-2-butanone and phenol again being significant (Table 16). Separation seen with PLS-DA supported further assessment of diagnostic utility (Figure 21). Diagnostic utility assessed with ROC analysis demonstrated an AUROC of 0.82 (95% CI 0.69-0.96), sensitivity 67% (95% CI 0.67-0.91) and specificity 93% (95% CI 0.86-1.0) (Figure 22).

**Table 16: VOCs found to be significantly different (p= ≤0.01), when assessing VOC abundance in those with no HCC and BCLC curative (early) HCC.**

| VOC | p value | False discovery rate |
|---|---|---|
| **3-methyl-2-butanone** | 0.001 | 0.07 |
| **o-cymene** | 0.001 | 0.07 |
| **D-limonene** | 0.001 | 0.07 |
| **Phenol** | 0.004 | 0.1 |
| **Styrene** | 0.01 | 0.2 |

False discovery rate applied in order to correct for multiple comparisons (q=≤0.05 for significance).

**Figure 21: PLS-DA based upon all VOC in all samples when assessing VOC abundance in those with no HCC and BCLC curative (early) HCC.**

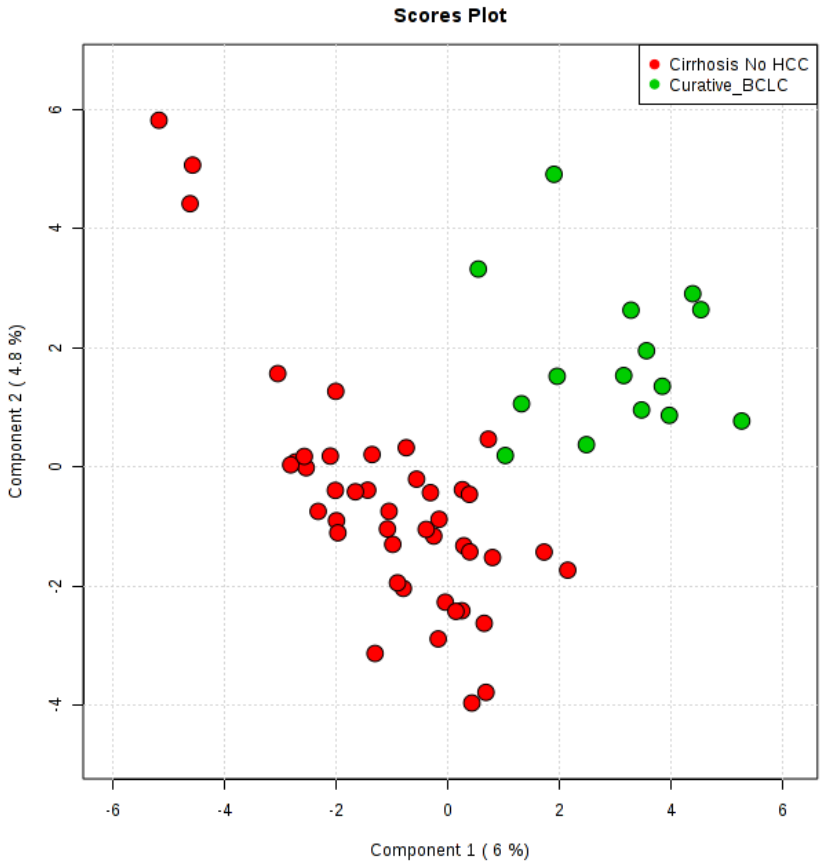
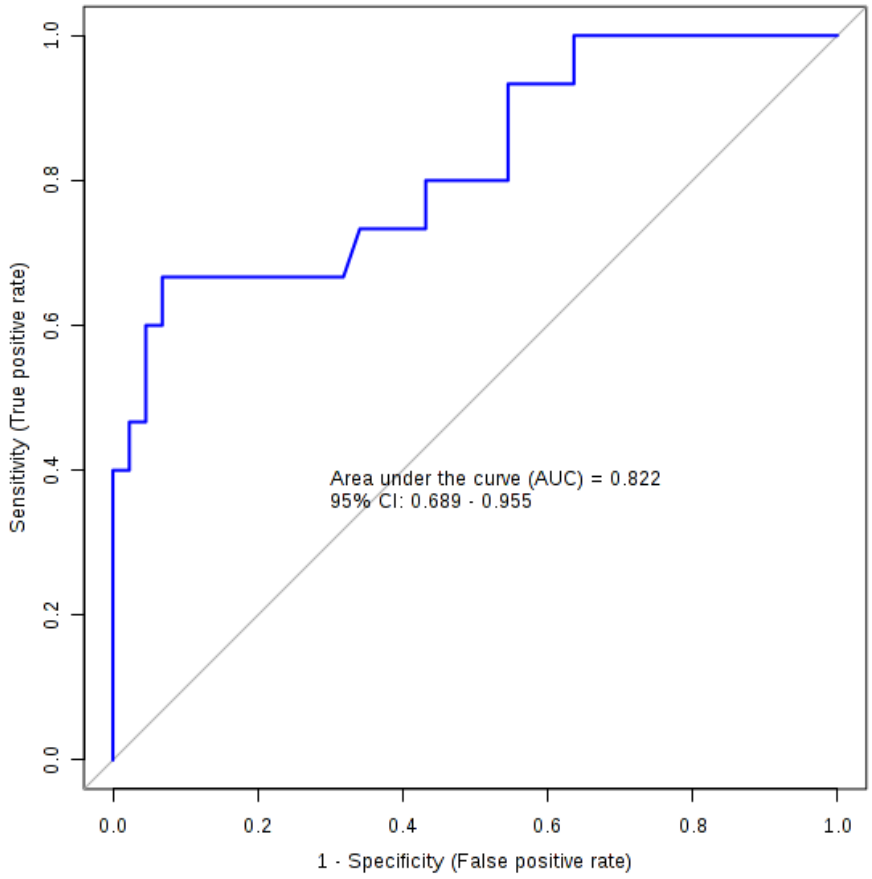


**Figure 22: ROC curve produced with logistic regression and 10 fold cross validation modelling, based upon the 5 VOC found to be significantly different for those with no HCC and BCLC curative (early) HCC.**



### 3.5 Discussion

This chapter reports the results of two pilot studies (smaller and larger) examining the utility of urinary VOCs as a biomarker for HCC. The smaller pilot study demonstrated differences according to the absences and presences of HCC. This significance did not persist after correction for false discovery and was not validated by the larger, Liverpool pilot study. The larger pilot study allowed for further stratification and modelling of VOC profiles, providing an accurate diagnosis of HCC and in particular for those who are treatment naive and staged earlier according to the BCLC stage. This has the potential to improve outcomes as early diagnosis has been shown to improve survival from HCC[127,129].

Currently in the UK, HCC surveillance is recommended for patients with underlying cirrhosis, EASL specifically recommend it for those who are Childs-Pugh A and B, Childs-Pugh C awaiting transplantation, non-cirrhotic HBV carrier with active hepatitis or family history of HCC and non-cirrhotic patients with chronic hepatitis C and advanced liver fibrosis[129]. For patients with preserved hepatic function, no evidence of portal hypertension and early HCC (a single lesion <5cm in diameter) resection has been shown to provide 5 year survival rates of 70%, with transplantation in such cases having five year survival rates up to 74%[325,326]. For those with early disease who are not deemed suitable for resection or transplantation, RFA with curative intent, has been shown to have a 5-year survival of approximately 37%[144]. The survival of those with later staged HCC is stark in comparison with the median survival of patients reported to be between 6 and 13.5 months depending on treatment mode[327]. Unfortunately less than 30% of patients are diagnosed at an early enough stage to make them eligible for resection or transplantation[328]. A meta-analysis form 2009 looked specifically at studies examining the impact of ultrasound based screening on detection of early HCC. Ultrasonography was demonstrated to be highly accurate for HCC at any stage, with pooled sensitivity of 94% and specificity of 94%. However, this was significantly lower when analysis was modified for the detection of early HCC, with the pooled sensitivity falling to 64%. Meta-regression analysis demonstrated a significantly higher sensitivity for the early HCC with an ultrasound scan every 6 months rather than with annual surveillance[144]. Results from the larger pilot study for the diagnosis of all treatment naive HCC was AUROC of 0.81 (95% CI 0.66-0.93), sensitivity 77% (95% CI 0.71-0.83) and specificity 75% (95% CI 0.71-0.79), with those defined as BCLC early HCC having an AUROC of 0.82 (95% CI 0.69-0.96), sensitivity 67% (95% CI 0.67-0.91) and specificity 93% (95% CI 0.86-1.0). Despite its limitations ultrasonography is currently the best screening tool for HCC. Our results suggest that urinary VOC analysis is superior to ultrasound scanning, particularly for the pertinent question of early HCC. Certainly, data from this meta-

analysis supports the use of 6 monthly screening, with a significant improvement in detection of early lesions when compared to 12 monthly (p= 0.001). We would therefore suggest urinary VOCs be assessed on a 6 monthly basis.

There may be a place for a two modality screening method combining both ultrasound scanning and VOC analysis. In previous clinical practice combined testing with AFP and ultrasonography was employed. The same meta-analysis demonstrated a non-statistically significant increase in sensitivity when AFP was used in addition to ultrasound scans[144]. This finding is consistent with current American Association of Liver Disease guidance which suggests that AFP is not an adequate screening tool, but has a role in the diagnosis of HCC when >200 ng/ml in the setting of a mass detected on imaging[329]. Moreover, normal AFP levels are present in as many as 30% of patients at time of diagnosis and usually remain low, even with advanced HCC[330]. This was even greater in our study, with 37/59 (62%) of those with HCC being AFP positive at diagnosis. We have demonstrated that VOCs are able to distinguish between HCC that is AFP negative and positive, along with those without HCC and AFP negative HCC.

Other potential surrogates to ultrasound screening are AFP-L3% and Des-gamma-carboxy prothrombin (DCP), neither of these reported biomarkers has been shown to be consistently superior to the demonstrated ability of urinary VOCs[228,232,233]. Dedicated trials examining the addition of urinary VOCs to ultrasound screening would need to be conducted and would need to include longer term survival data.

Four compounds were recurrently highlighted to be important for the identification of HCC, with all four being significant for the detection of early HCC. These were 3-methyl-2-butanone, styrene, phenol and D-limonene. 3-methyl-2-butanone was most abundant in those without HCC, with the remaining three compounds being most abundant in those with early HCC. Styrene is metabolised in the liver by the cytochrome P450-mediated monooxygenase system. A number of genetic studies have suggested that individual susceptibility to cancer may be determined by an inherited genetic predisposition associated with the polymorphisms of genes encoding the enzymes involved in such metabolism, such polymorphism as CYP1A1 MspI[331]. The products of metabolism are excreted in the urine. The National Toxicology Programme in the USA report that styrene is reasonably anticipated to be a human carcinogen based on limited evidence of carcinogenicity from studies in humans, sufficient evidence of carcinogenicity from studies in experimental animals, and supporting data on mechanisms of carcinogenesis[332]. An accumulation of styrene as a result of impaired liver function may be one reason for the increase seen in our cohort. Most of the genetic damage associated with styrene exposure is thought to be due to styrene-7,8-oxide, a by-product of its

metabolism[333]. D-limonene, although at high doses have been shown to cause renal cancer in male rats, is considered by most researchers to be a potential chemo-preventive agent and has been declared via international consensus to have no carcinogenic potential. In fact it has been suggested as having an efficacious chemotherapeutic agent for human malignancies[334]. Its greater abundance in those with HCC may therefore be related to altered metabolism and accumulation, rather than a direct causative effect. Phenol in contrast to this, is recognised as being carcinogenic to humans, namely in colorectal cancer as part of bacterial fermentation[287]. Despite this, the International Agency for Research on Cancer (IARC) classified phenol as a category 3 carcinogen, i.e. not classifiable for its carcinogenicity to humans[335]. It is mostly metabolised in the liver and excreted in the urine. In the liver phenol is metabolised by sulfotransferases, specifically phenol sulfotransferase 1[336]. This has been shown to be down-regulated in HCC and this potentially explains the accumulation of phenol in the urine of HCC patients[331]. Once again the increased phenol may be accumulation due to altered metabolism rather than a direct causative effect.

Qin *et al* assessed VOCs emitted in the breath of 30 HCC patients who were infected with HBV and cirrhosis and from 27 cirrhotic patients without HCC, together with 36 healthy persons, both taken as controls. 3-hydroxy-2- butanone, styrene, and decane were identified as having the most potential as biomarkers for HCC. Only 3-hydroxy-2- butanone (AUROC 0.75 [0.61-0.87], sensitivity 70% and specificity 70.4%) and styrene (AUROC 0.69 [0.544-0.829], sensitivity 66.7% and specificity 70.4%) were able to distinguish between HCC and cirrhotics without HCC. They did not find any difference in those that were deemed to be AFP positive or negative[337]. In our pilot study styrene performed strongly in the diagnosis of all HCC, treatment naive and early stage HCC. Mochalski *et al* performed VOC analysis of HCC cell culture in an attempt to assess the metabolites of the cells. They characterised 9 VOCs that were categorised as HCC uptake and 12 that were released by the HCC cells[338]. None of these 21 VOCs was identified as being significant in our analysis. A further study compared VOCs emitted from 31 HCC tissue and corresponding non-tumour liver tissue from the same liver. HCC was characterized by approximately two-fold depletion of glucose, glycerol 3- and 2-phosphate, malate, alanine, myo-inositol and linoleic acid[339]. Once again these compounds were not found to be significant in our study.

Using the larger pilot study data, validation of the differences seen in the smaller pilot samples was attempted and failed. There are a number of potential reasons why this may have occurred. Firstly, the volumes of urine and the headspace vials used were different between the two groups. The smaller pilot study used 400 ul of urine in a 2ml headspace vial, whilst the pilot study used 4 ml of urine in a 10 ml headspace vial. Previous work, together with the increased mean VOC yield in the

pilot study, had demonstrated that 4 ml was the optimal volume to use[317]. VOCs may therefore not have been identified by the sub-optimal methodology employed in the smaller pilot study. In some instances, samples from the smaller pilot study were more than 3 years old, whilst the maximum age of the pilot study samples was 6 months. Therefore, there is likely to have been a greater degree of VOC loss and sample degradation owing to the difference in storage duration. Before taking custody of the smaller pilot samples they had been stored in soft plastic vials, such plasticizers are recognised contaminators of GCMS results. The pilot study samples were not exposed to such plasticizers. A further point of note is the geographical location of recruitment and thus the potential for large variation in the ethnicity of each cohort. The smaller pilot cohort were all recruited from the Queen Elizabeth Hospital, Birmingham UK, whilst the larger pilot study patients were recruited from Liverpool, UK. This point is demonstrated in by results from the 2011 census (Table 17). Dietary differences will be seen between geographical location and ethnicity, therefore making the range of VOCs identified in urine different. Any future studies assessing urinary VOCs should aim to recruit patients in multiple geographic locations to correct for this factor.

**Table 17: Results from 2011 UK census showing the difference in population size and ethnic composition of Liverpool and Birmingham**[340]**.**

| | | **White** | | **Asian or Asian British** | | | | **Black or Black British** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **White British & Irish** | | **Indian** | | **Pakistani** | | **Black African** | | **Black Caribbean** | |
| **City** | Population | n= | % | n= | % | n= | % | n= | % | n= | % |
| Liverpool | 466,415 | 402,399 | 86.3 | 4,915 | 1.1 | 1,999 | 0.4 | 8,490 | 1.8 | 1,467 | 0.3 |
| Birmingham | 1,073,045 | 592,646 | 55.2 | 64,621 | 6.0 | 144,627 | 13.5 | 29,991 | 2.8 | 47,641 | 4.4 |

False discovery rate testing was performed in Metaboanalyst. The numerical figure generated represents a percentage, indicating the expected false positives among all features predicted to be significant. The significance of the VOCs that were initially identified by the smaller Birmingham pilot study did not persist after a false discovery rate correction was applied. This is likely to be because the smaller numbers made the study under-powered. It may also explain why these differences were not seen in the larger, adequately powered, Liverpool study. Across the different sub-group analyses performed for the Liverpool pilot study a number of compounds achieved, or were close to, a significant false discovery rate correction. These were styrene, D-limonene, phenol and 3-methyl-2-butanone. Again some of this variation may be related to smaller numbers of patients in elements of the sub-group analysis. These compounds featured in the biomarker modelling that assessed diagnostic utility. Combining inclusion of VOCs with significant false

discovery rate (q=≤0.05) and performing logisitic regression together with 10 fold cross validation can limit overfitting, a common problem when identifying biomarkers[341].

### 3.6 Conclusion

Analysis of VOCs emitted from urine has the potential to discriminate those cirrhotic patients who have developed HCC. Moreover, it can most accurately identify those who are treatment naive and who have early stage HCC. This has the potential to improve long term survival as curative treatment is more likely if HCC is diagnosed at an earlier stage. It appears that 4ml of urine and the use of a 10 ml headspace vial is superior to lower volumes and a 2ml headspace vial and that geographic differences may exist in the VOCs emitted from urine.

# Chapter 4

## An investigation of volatile organic compounds found in the headspace gas of cultured *Fusobacterium nucleatum* and *Campylobacter showae*

### 4.1 Introduction

The human intestinal microbiota contains a plethora of diverse microbial species, wherein certain bacteria, considered to be driver bacteria with carcinogenic features, contribute directly toward colonic epithelial cell damage and genetic mutations to initiate colorectal carcinogenesis: *Campylobacter* spp are such bacteria. However, some bacteria, in particular, *Fusobacterium nucleatum*, which is otherwise a normal resident of the oral microflora and a relatively poor colonizer of the healthy gut, have also been considered to play a role in the development of colorectal cancer. The two species have been implicated in the driver-passenger model of colorectal cancer [342,343].

#### 4.1.1 Aim

- To characterise the VOCs emitted from a culture medium containing *Fusobacterium nucleatum.*
- To characterise the VOCs emitted from a culture of *Campylobacter showae*
- To assess the changes in VOCs emitted from co-cultures of *Fusobacterium nucleatum* and *Campylobacter showae*.

#### 4.1.2 Method

Bacterial cultures were grown by a collaborator, Dr G Hold, in the School of Public Health, Harvard University, Boston, USA. They consisted of *F. nucleatum* alone, *C. showae* alone, 30 min co-culture and a 24 hour co-culture. The *F. nucleatum* and *C. showae* alone samples were cultured for 24 hours. A SPME fibre was then exposed to the headspace of the culture vials. The fibres were then shipped to our laboratory in Liverpool, UK. GCMS analysis was subsequently performed and AMDIS used for VOC identification and quantification. Statistical analysis used the programme R and a dedicated in-house pipeline.

### 4.2 Results

35 fibres were exposed to the headspace gas and shipped from USA to Liverpool. One fibre that was exposed to the *C. showae* culture failed to capture any volatiles and so was not included in the analysis. Therefore, 34 fibres underwent full analysis (Table 18).

**Table 18: Description of the headspace vial contents and the number of fibres exposed to the headspace gas.**

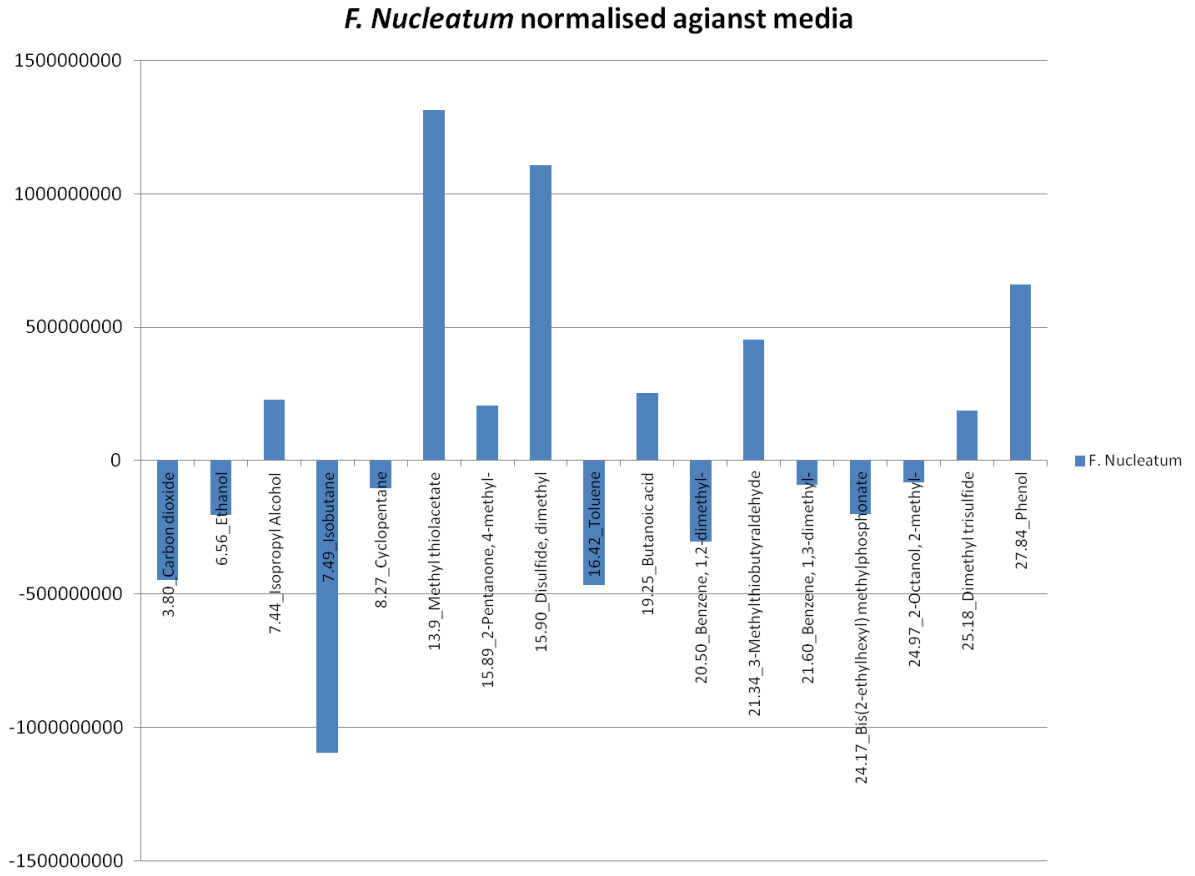
| Headspace vial content | Number of fibres exposed |
|---|---|
| **Media alone** | 4 |
| ***F. nucleatum* and media** | 6 |
| ***C. showae* (strain MSG) and media** | 9 |
| **Overnight co-culture of *F. nucleatum* and *C. showae* (24 hours)** | 8 |
| **30 minute co-culture of *F. nucleatum* and *C. showae*** | 8 |

### 4.2.1 Normalisation against media

### 4.2.2 *F. nucleatum* normalised against media

Using the statistical package R, volatiles identified in the *F. nucleatum* cultures were normalised against media. This enabled the VOCs to be classified as "produced" if there was a positive fold change and "utilised", if the fold change was negative (Figure 23).

**Figure 23: Bar chart showing those compounds with the largest fold changes when *F. nucleatum* was normalised against media.**



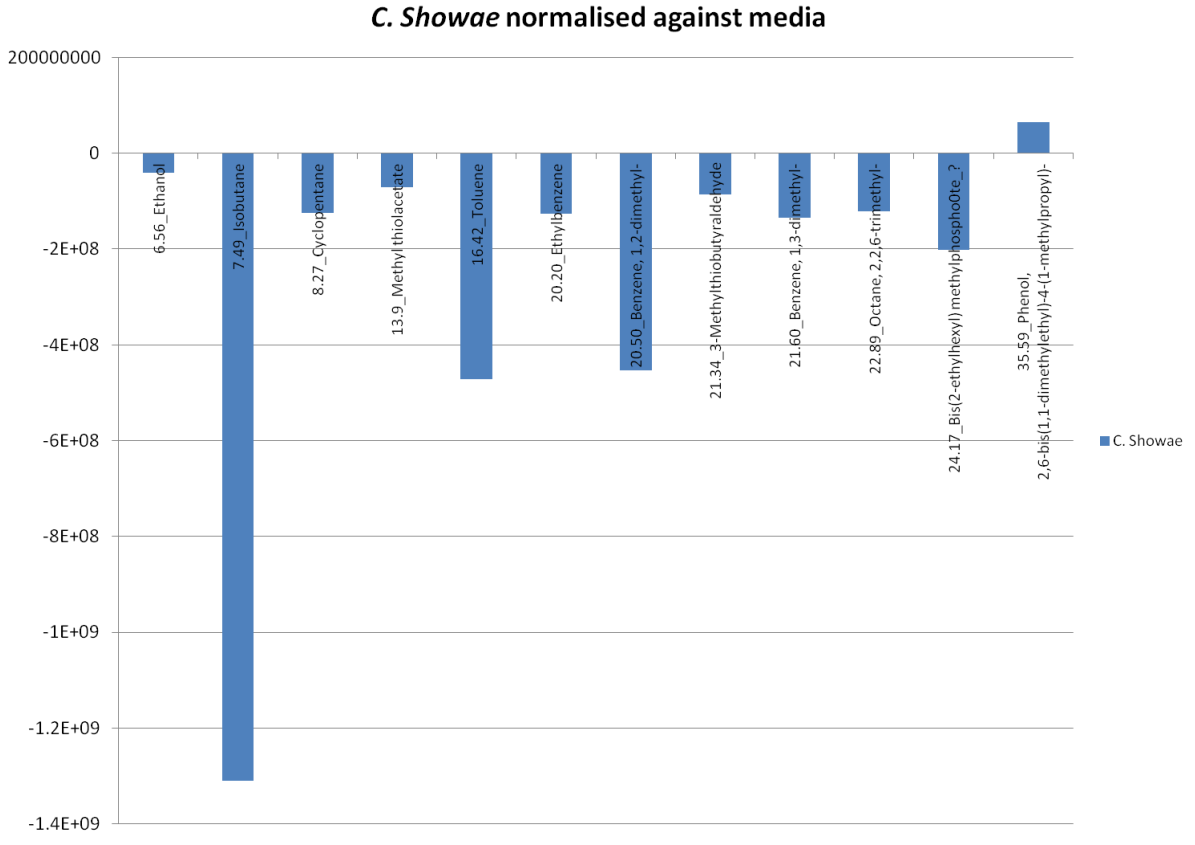
A negative value suggests utilisation of the compound by *F. nucleatum*, whilst a positive value is suggestive of production. The number prior to the compound name is the retention time. Y axis represents relative ion abundance.

There was significant production of methylthiolacetate, butanoic acid, dimethyl disulfide, dimethyl trisulfide and phenol by the *F. nucleatum*. Isobutane and toluene were associated with the largest negative fold changes.

### 4.2.3 *C. showae* normalised against media

The same analysis was then performed from data from *C. showae* and again produced the labels of "produced" and "utilised". When compared to *F. nucleatum* there is clearly more utilisation from the media than production, in the *C. showae* culture (Figure 24).

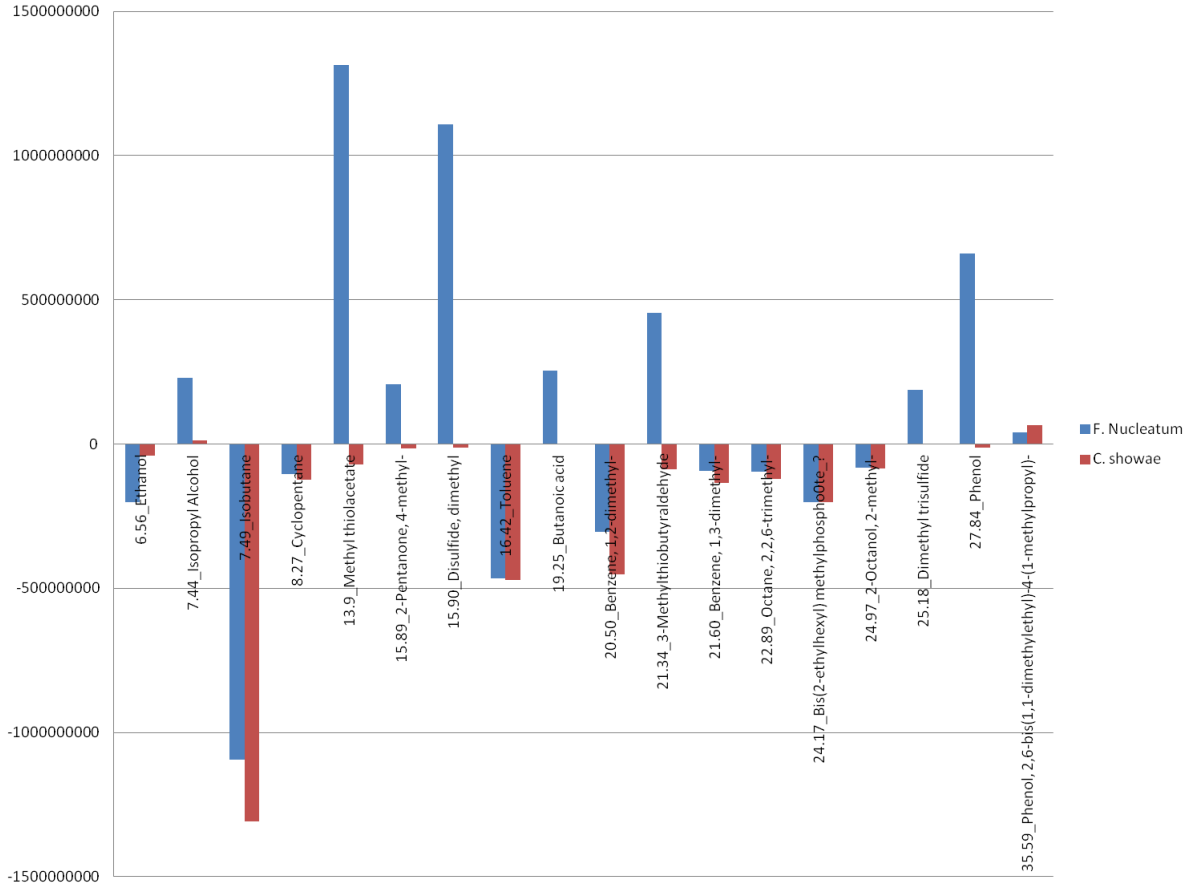
**Figure 24: Bar chart showing those compounds with the largest fold changes when *C. showae* is normalised against media.**




The number prior to the compound name is the retention time. Y axis represents relative ion abundance.

When these two bar charts are combined it is clear that *F. nucleatum* is a net "producer" of volatile organic compounds whilst *C. showae* is a net "utiliser" (Figure 25), with the utilisation pattern of the two showing a number of similarities.

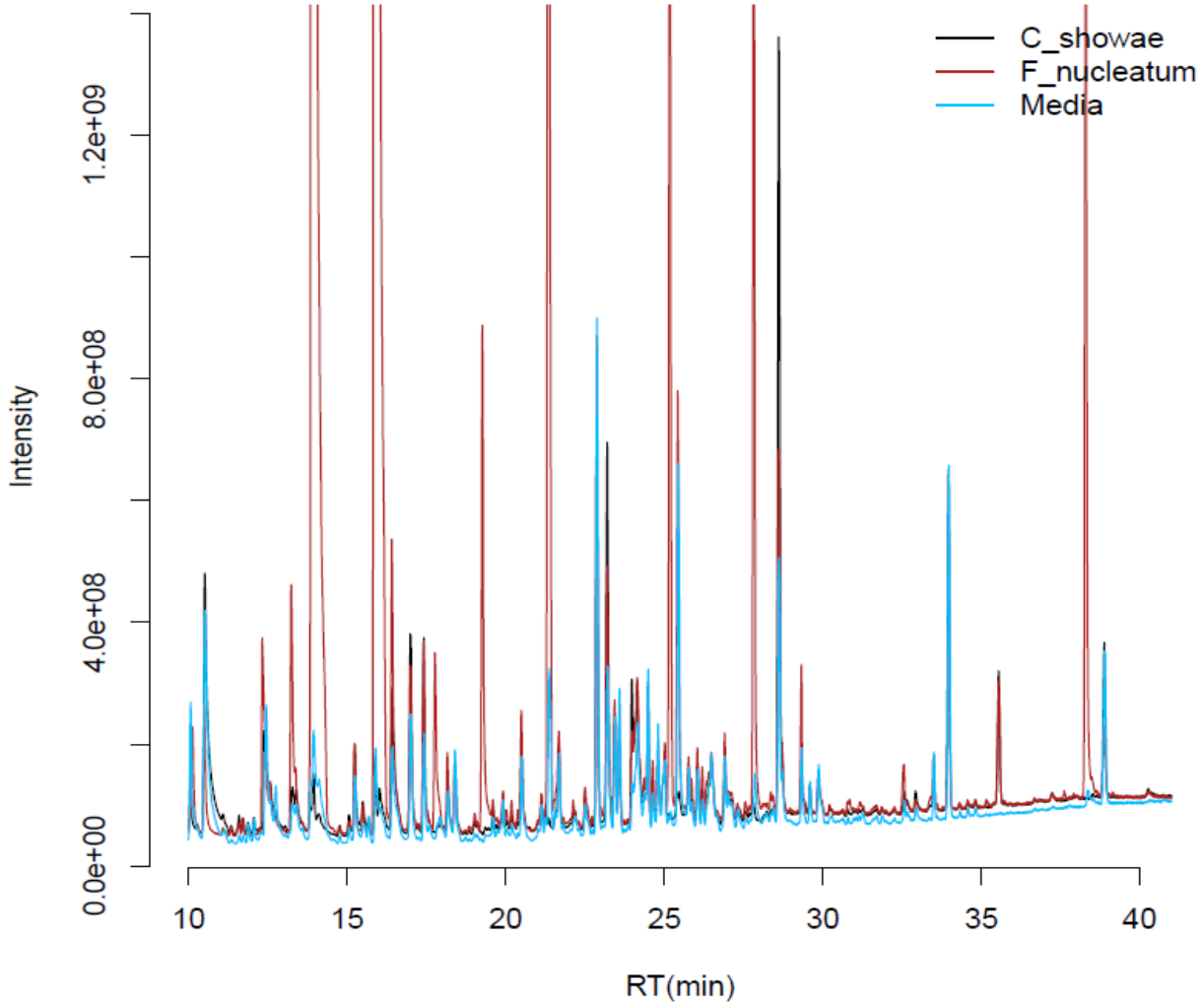
**Figure 25: Bar chart demonstrating the relative abundance of those volatile organic compounds with the largest fold changes when *F. nucleatum* and *C. showae* are normalised against the media.**




The number prior to the compound name is the retention time. Y axis represents relative ion abundance.

This point is further emphasised when the chromatograms are compared by overlaying them (Figure 26); there is a clear dominance, in terms of qualitative and quantitative production, in the *F. nucleatum* cultures.

**Figure 26: Overlay representative chromatogram derived from F. nucleatum, C. showae and media samples.**

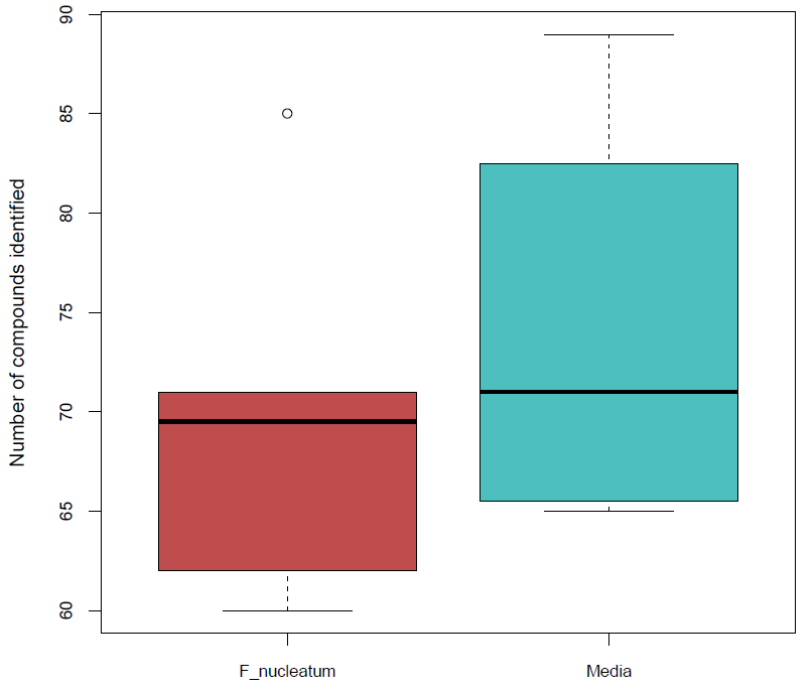


### 4.3 Direct Volatile Organic Compound Analysis

#### 4.3.1 Direct Comparison of *F. nucleatum* and media alone

No normalisation took place for this element of the analysis. There was no significant difference in the number of VOCs identified between the *F. nucleatum* and media alone groups (p=0.5) (Figure 27).

**Figure 27: Boxplot demonstrating the number of compounds identified in the *F. nucleatum* group and media alone group.**



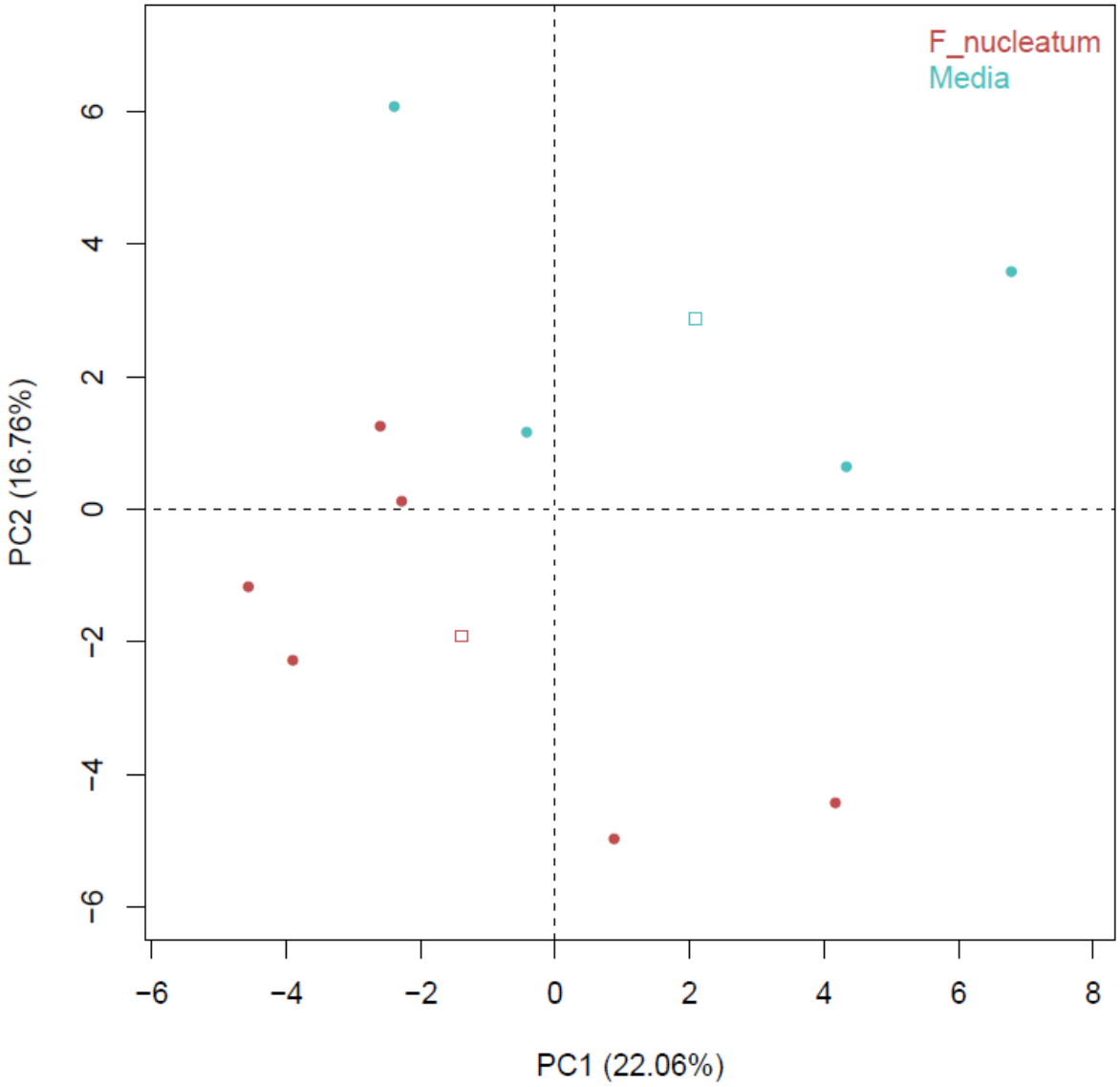

Categorical analysis looking at prevalence of a VOCs showed 7 to be significant different (p= <0.05) (Table 19).

**Table 19: Volatile organic compounds identified to be significantly different following assessment of frequency.**

| VOC | Present *F. nucleatum* | Absent *F. nucleatum* | Present media | Absent media | Fisher's exact test |
|---|---|---|---|---|---|
| **2-Pentanone** | 0 | 6 | 4 | 0 | 0.004 |
| **Butanoic acid** | 6 | 0 | 0 | 4 | 0.004 |
| **Methylcyclohexane** | 0 | 6 | 3 | 1 | 0.033 |
| **Tetrachloroethylene** | 0 | 6 | 3 | 1 | 0.033 |
| **4-methyl-2-pentanone** | 1 | 5 | 4 | 0 | 0.047 |
| **2,4,6-trimethyldecane** | 1 | 5 | 4 | 0 | 0.047 |
| **1-Propanol** | 5 | 1 | 0 | 4 | 0.047 |

*F. nucleatum* culture n=6 and media culture n=4

**Figure 28: Principal component analysis comparing *F. nucleatum* and media alone, based on prevalence.**

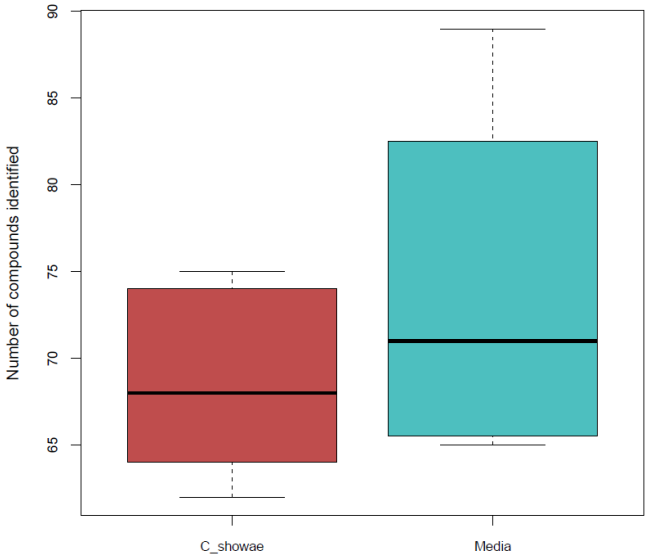



Coloured squares represent a centroid samples calculated from the mean of PC1 and PC2.

### 4.3.2 Direct Comparison of *C. showae* and media alone

There was no difference in the number of VOCs identified between the two groups (Figure 29).

**Figure 29: Boxplot demonstrating the number of compounds identified in the *C. showae* group and media alone group.**




Analysis of frequency and abundance did not show any significant differences.

### 4.3.3 Comparison of *F. nucleatum, C. showae* and media

Principal component analysis showed clustering of *C. showae* and, to a lesser extent, *F. nucleatum* with the media alone lying between the two (Figure 31). There was no difference in the number of VOCs identified between the 3 groups (Figure 30). Further assessment of prevalence was in line with prior analysis (Table 19). Butanoic acid, phenol, methylthiolacetate and sulphides were predominant in the *F. nucleatum* samples across all elements of the analysis (Tbale 20).

**Figure 30: Boxplot demonstrating the number of VOCs identified in the 3 groups, *C. showae, F. nucleatum* and media.**

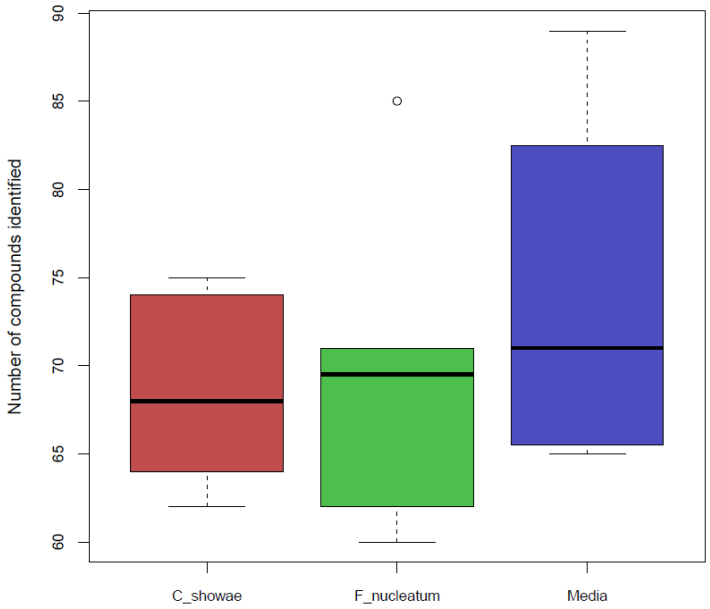
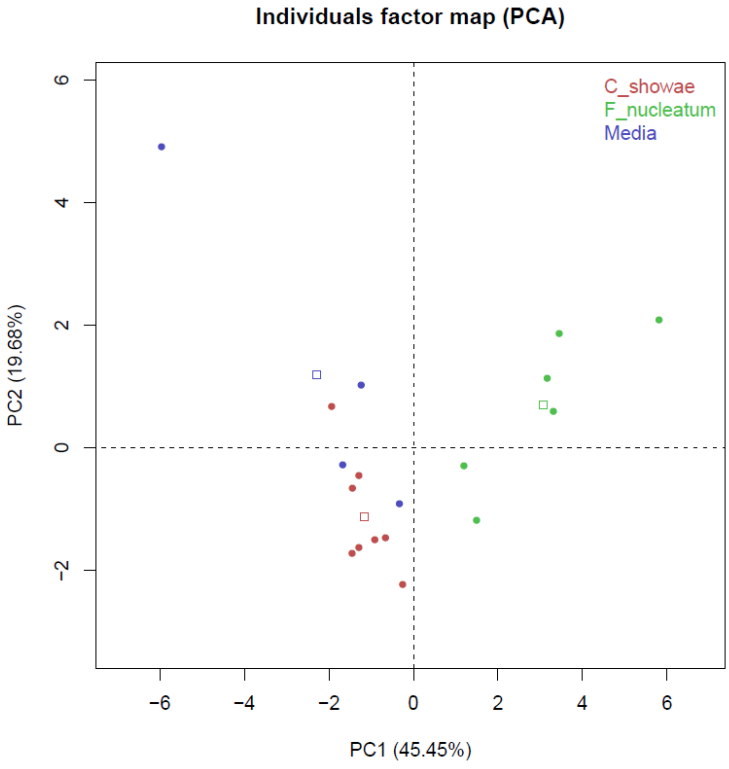



**Figure 31: Principal component analysis comparing based upon relative abundance comparing *F. nucleatum*, *C. showae* and media.**




Coloured squares represent a centroid samples calculated from the mean of PC1 and PC2

**Table 20: Table demonstrating the presence and absence of VOCs in *C. showae*, *F. nucleatum* and media that achieved a p value of <0.05.**

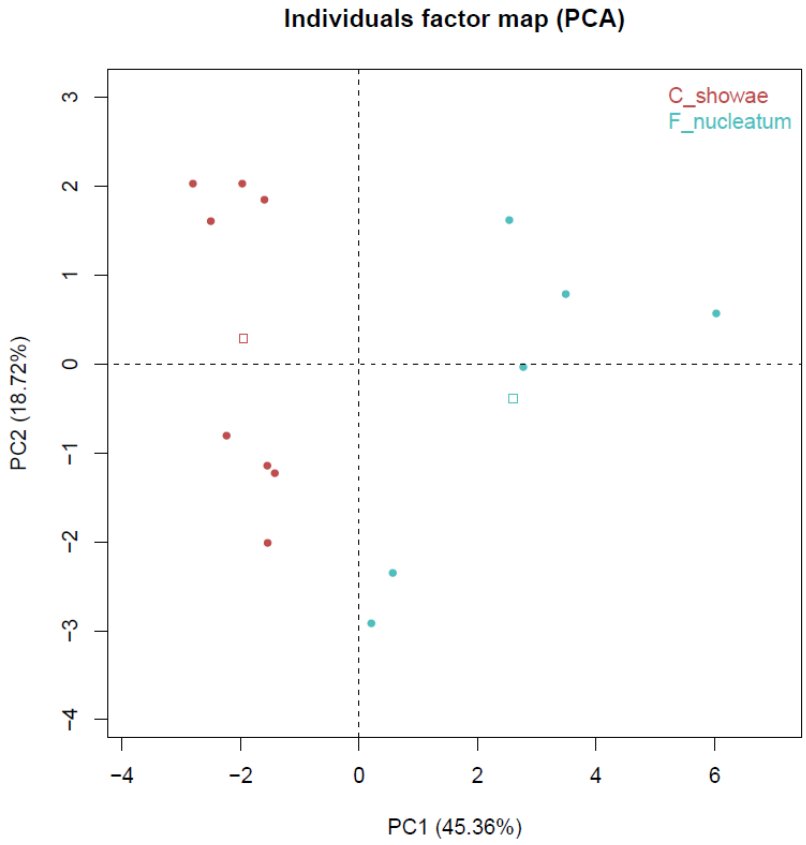
| | Present *C. showae* | Absent *C. showae* | Present *F. nucleatum* | Absent *F. nucleatum* | Present media | Absent media | Fishers exact test |
|---|---|---|---|---|---|---|---|
| 2-pentanone | 8 | 0 | 0 | 6 | 4 | 0 | <0.0001 |
| Butanoic acid | 0 | 8 | 6 | 0 | 0 | 4 | <0.0001 |
| 4-methyl-2-pentanone | 8 | 0 | 1 | 5 | 4 | 0 | 0.001 |
| 1-propanol | 1 | 7 | 5 | 1 | 0 | 4 | 0.01 |
| 2,6-bis(1,1-dimethylethyl)-4-(1-methylpropyl)-phenol | 8 | 0 | 5 | 1 | 1 | 3 | 0.01 |
| Methyl-cyclohexane, | 5 | 3 | 0 | 6 | 3 | 1 | 0.02 |
| Dimethyldisulfide, | 0 | 8 | 4 | 2 | 2 | 2 | 0.02 |
| Methylthiolacetate | 0 | 8 | 4 | 2 | 1 | 3 | 0.02 |
| 3-Methylthiobutyraldehyde | 0 | 8 | 4 | 2 | 1 | 3 | 0.02 |
| Phenol | 0 | 8 | 4 | 2 | 1 | 3 | 0.02 |
| Dimethyl trisulfide | 0 | 8 | 3 | 3 | 0 | 4 | 0.03 |
| 2,4,6-trimethyl-decane | 5 | 3 | 1 | 5 | 4 | 0 | 0.04 |
| Tetrachloroethylene | 4 | 4 | 0 | 6 | 3 | 1 | 0.04 |

Fisher t test conducted using presence and absence in *C. showae* and *F. nucleatum* cultures.

### 4.3.4 Direct comparison of *C. showae* and *F. nucleatum*

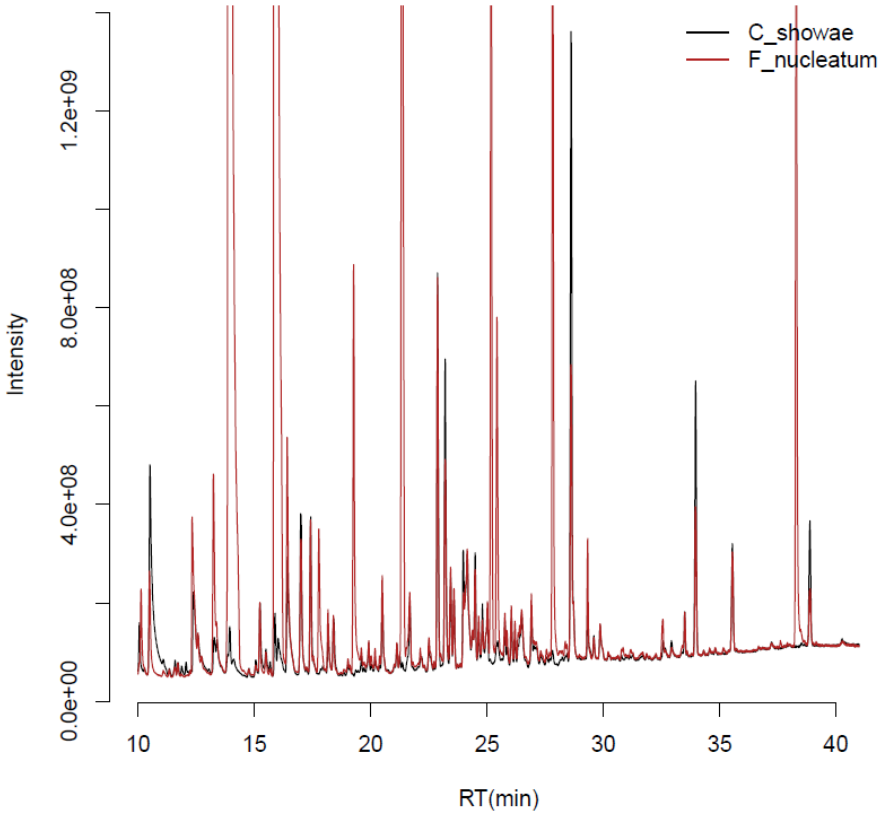
The separation of the *F. nucleatum* from *C. showae* was shown by PCA when *F. nucleatum* and *C. showae* were compared without including media (Figure 32).

**Figure 32: Principal component analysis comparing based upon relative abundance comparing *F. nucleatum* and *C. showae*.**




Coloured squares represent a centroid samples calculated from the mean of PC1 and PC2

**Figure 33: Overlay representative chromatograms derived from *F. nucleatum* and *C. showae***



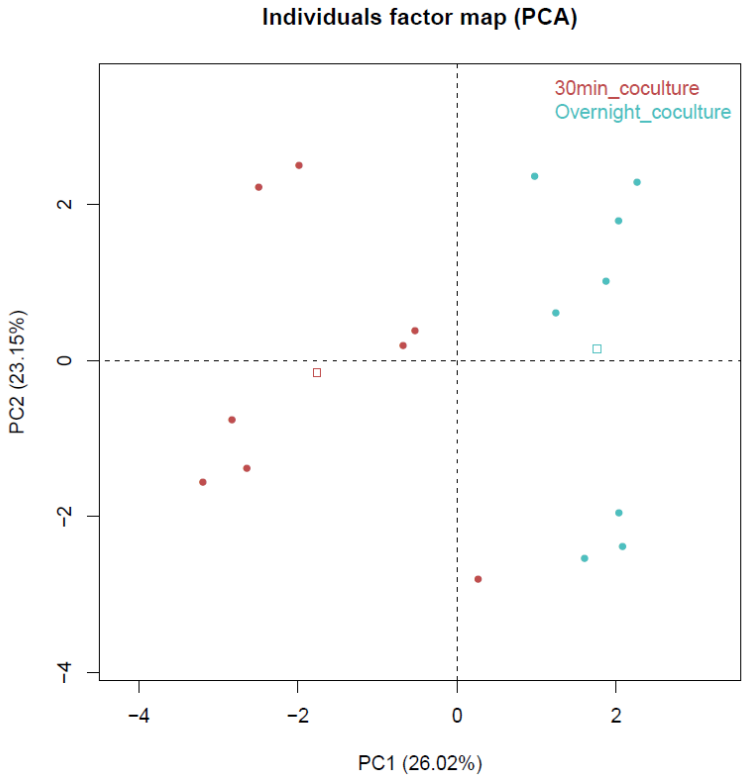

The same differences were identified: there were significant differences in butanoic acid, phenol, methylthiolacetate and sulphides.

**4.4 Analysis of co-cultures to assess the potential impact of *C. showae* on *F. nucleatum***

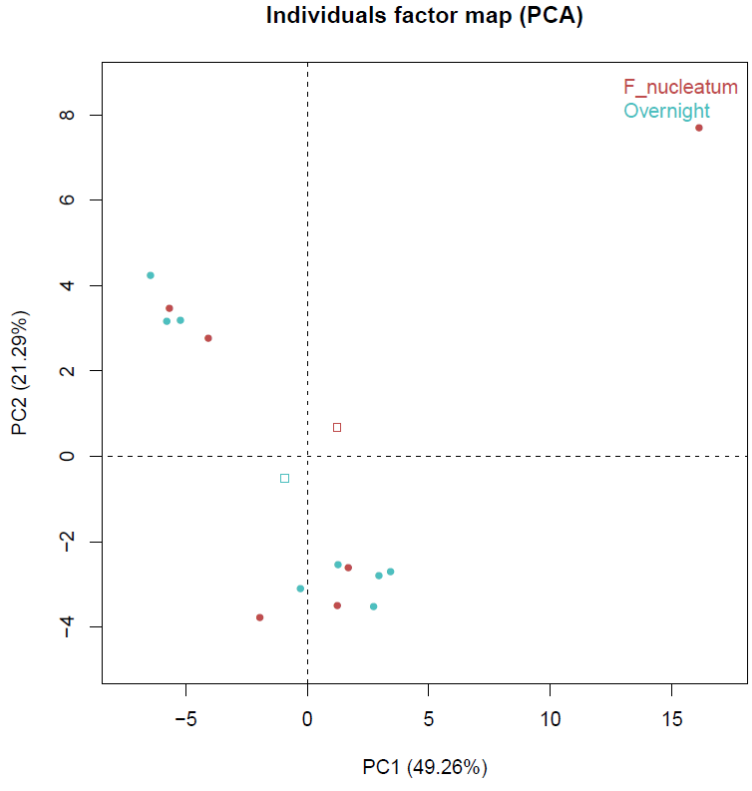
DNA sequencing data, by others, has shown that, after 24 hours of co-culture, there is a significant over-representation of *F. nucleatum*, even with an initial 6:1 ratio of *C. showae* to *F. nucleatum*. There was no difference in the number of VOCs identified in the *F. nucleatum*, 30 minute and 24 hour co-cultures(Hold *et al* 2015, unpublished).

Principal component analysis suggested a difference when comparing 30 min co-culture and 24 hour co-culture (Figure 34a). When the same analysis is applied to *F. nucleatum* culture and 24 hour co-culture, this degree of separation is lost and the samples appear to homogenise (Figure 34b).

**Figure 34a and b: (a) Principal component analysis comparing based upon relative abundance comparing 30 minute and 24 hour co-culture. (b) Principal component analysis comparing based upon relative abundance comparing *F. nucleatum* and 24 hour co-culture.**




(a)




(b)

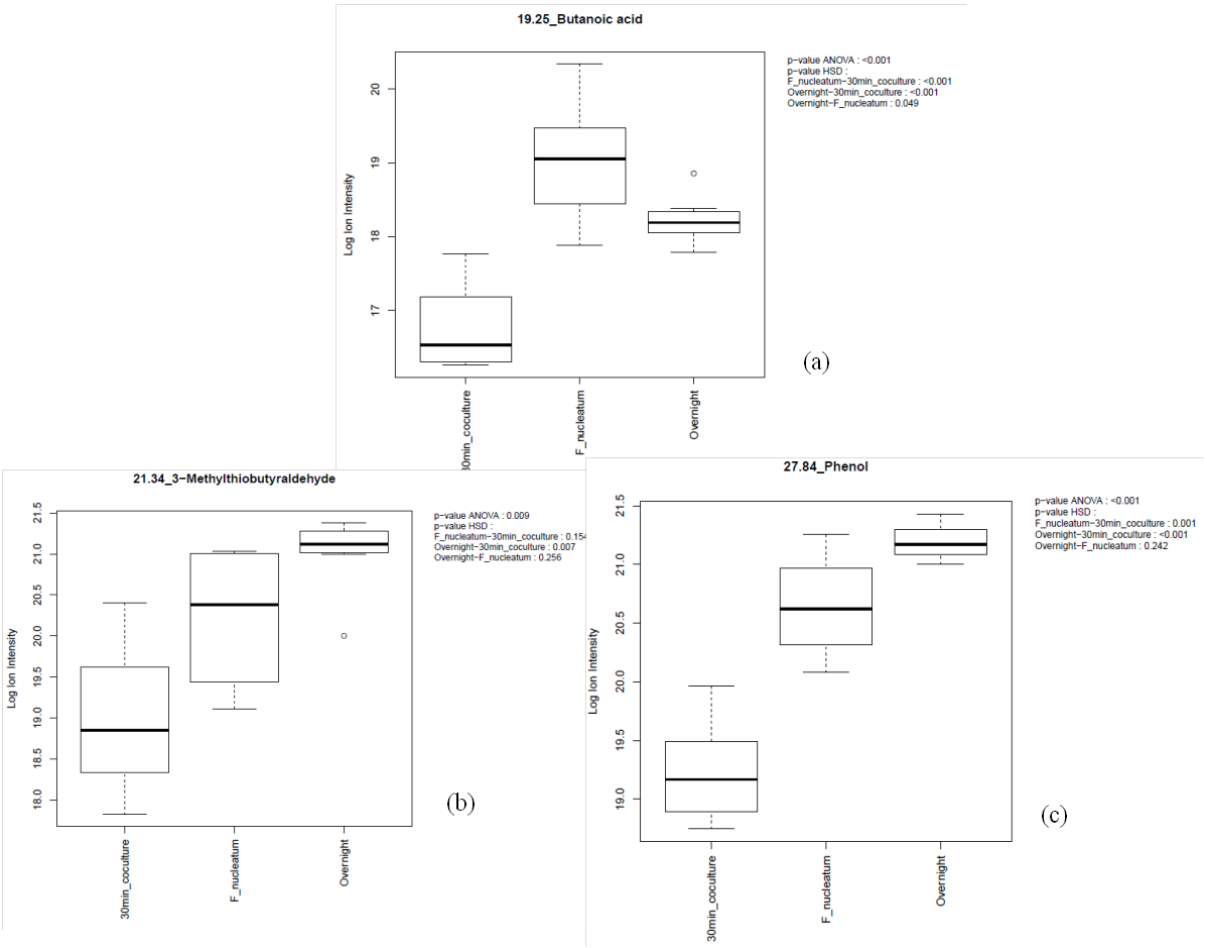
Coloured squares represent a centroid samples calculated from the mean of PC1 and PC2

Butanoic acid, 3-methylthiobutyraldehyde, disulfide dimethyl and methylthioacetate are all most abundant in the *F. nucleatum* alone samples, and are lowest in the 30 min co-culture and recovering in the 24 hour co-culture. Phenol is most abundant in the 24 hour co-culture and least in the 30 min co-culture (Figure 36). Comparison of abundance using ANOVA was performed across the 3 groups and identified 3 VOCs to be significantly different (Table 21). This data is also presented as box plots in Figure 35a, b and c.

**Table 21: VOCs identified as significantly different by ANOVA performed on *F. nucleatum* alone, 30 minute co-culture and 24 hour co-culture.**

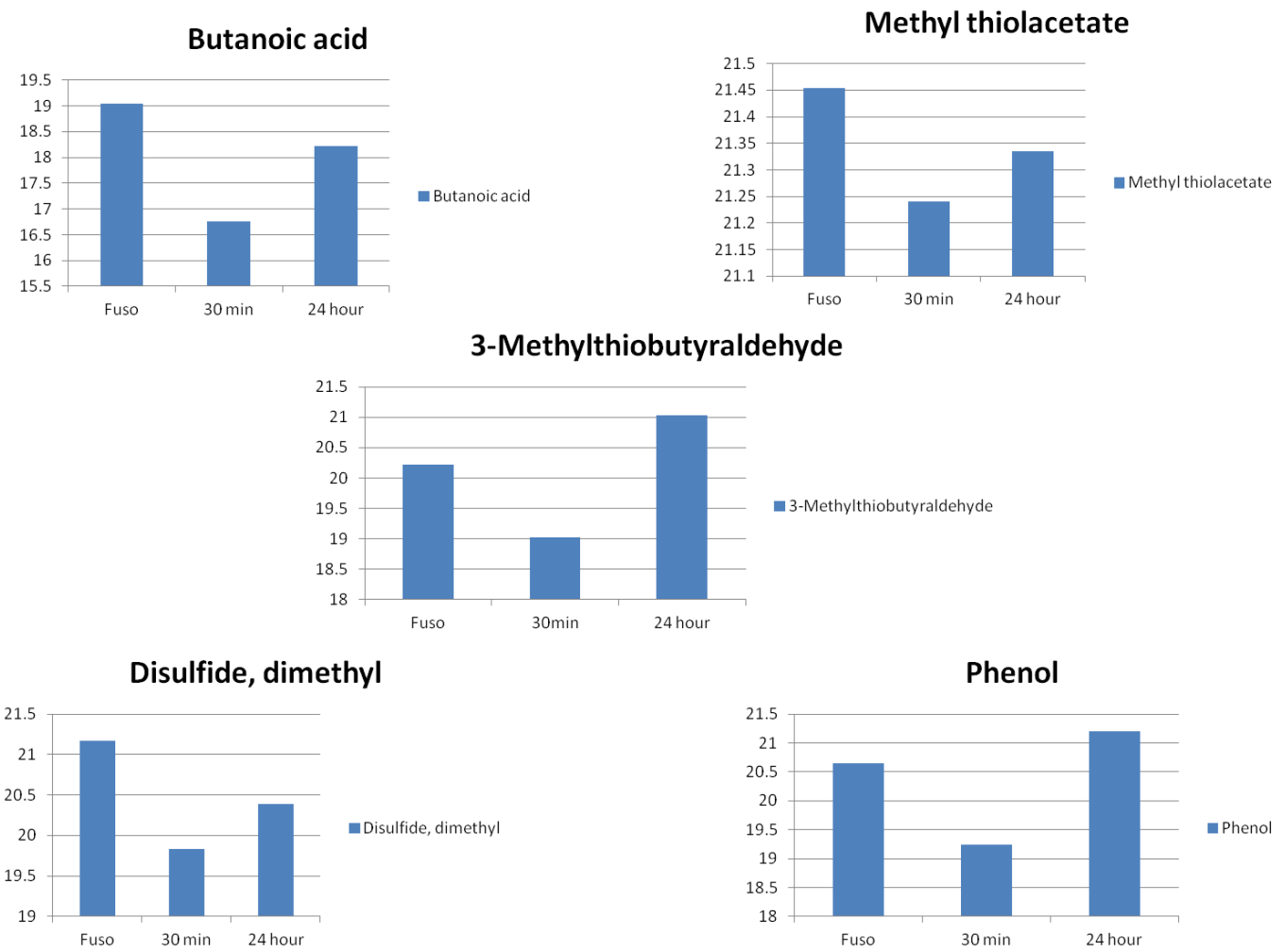
| | Mean log transformed abundance *F. nucleatum* alone | Mean log transformed abundance 30 min co-culture | Mean log transformed abundance 24 hour co-culture | ANOVA p value |
|---|---|---|---|---|
| **Butanoic acid** | 19.04 | 16.75 | 18.22 | <0.0001 |
| **Phenol** | 20.64 | 19.23 | 21.19 | 0.0001 |
| **3-Methylthiobutyraldehyde** | 20.22 | 19.02 | 21.02 | 0.008 |

**Figure 35a, b and c: Box plots representing the relative ion abundance of the 3 VOCs identified as significantly different by ANOVA. The features that provide the significance are provided as text to the right of each plot.**




Key- Overnight represents the 24 hour co-culture of *F. nucleatum* and *C. showae*, 30 min represents co-culture of *F. nucleatum* and *C. showae* samples after 30 mins, *F. nucleatum* represent isolated *F. nucleatum* culture.

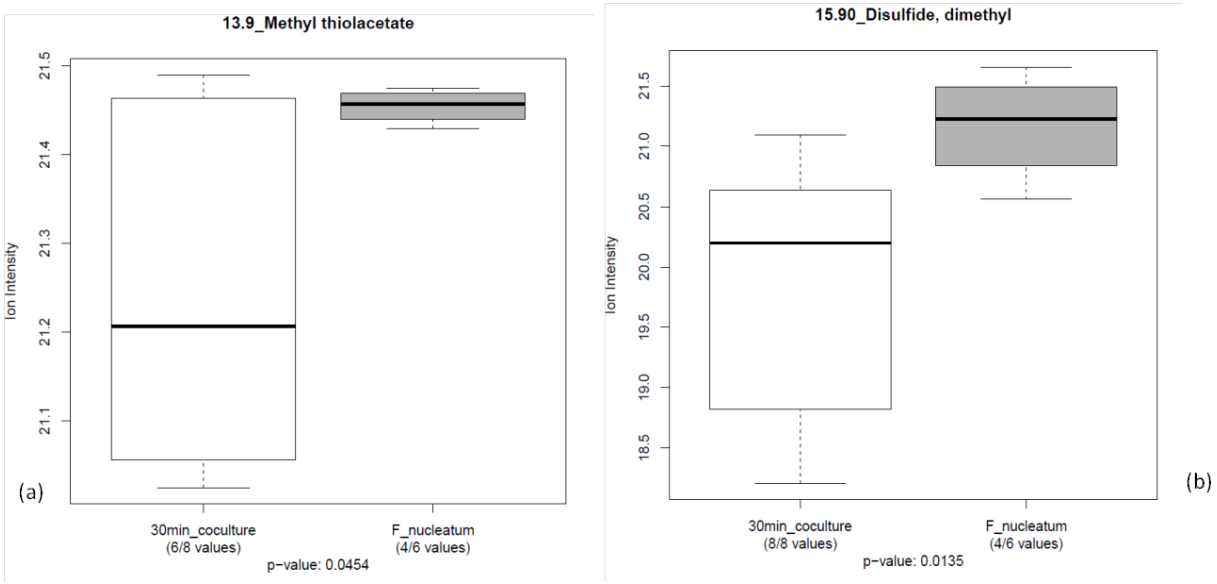
**Figure 36: Bar charts for those VOCs found to have significantly different abundances when comparing 24 hour *F. nucleatum* alone, 30 minute co-culture and 24 hour co-culture.**




Y axis presents log transformed relative ion abundance, Fuso= Fusobacterium nucleatum culture, 30 min and 24 hour= duration of co-culture.

When comparing the abundance of *F. nucleatum* alone and the 30 minute co-culture, two additional VOCs were significantly different: disulfide dimethyl (p=0.013) and methylthioacetate (p=0.04) (Figure 37a and b).

**Figure 37a and b: Box plots for methylthiolacetate and dimethyldisulfide.**




Y axis is log transformed relative ion abundance.

When comparing the *F. nucleatum* and 24 hour co-culture for prevalence and abundance, no significant differences were found.

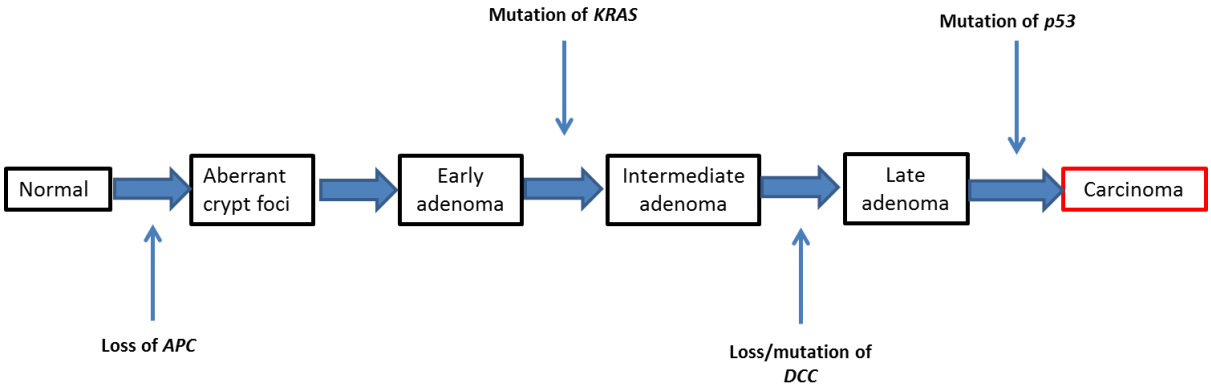
**4.5 Discussion**

There is a clear difference in the metabolic activity of *F. nucleatum* and *C. showae* as shown by the VOCs identified during this study. When the VOCs are normalised against the VOCs within the media there is clear evidence that *F. nucleatum* is a net producer of compounds whilst *C. showae* is a utiliser. This is mirrored when there is comparison across the VOCs produced by *F. nucleatum* alone, the 30 minute co-culture and the 24 hour co-culture. Throughout the different elements of the analysis the same VOCs were significant, namely, butanoic acid, phenol, methylthiolacetate, sulphides and 3-methylthiobutyraldehyde. These are more prominent and/or abundant when the representation of *F. nucleatum* is greatest.

It has been suggested that worldwide approximately 20% of cancers are related to infectious agents, for example human papilloma virus and cervical cancer, along with *Helicobacter pylori* and gastric

cancer [344]. Of the estimated 3.7 x $10^{30}$ microbes living on earth, only 10 are designated by the International Agency for Cancer Research (IACR) as carcinogenic to humans[345]. Despite this there are also a number of other bacteria with robust data supporting their role in human carcinogenesis, with a number being implicated in the pathogenesis of colorectal cancer. Bacteria constitute about 90% of all the cells in the human body and it has been estimated that bacterial genes outnumber human genes by two orders of magnitude or more[346]. The bacterial density of the large bowel (~$10^8$ cells per ml) is much greater than that of the small bowel (~$10^2$ cells per ml), this is paralleled by an approximate 12-fold increase in cancer risk for the large bowel compared to the small bowel [343].

**Figure 38: Adenoma-carcinoma sequence with associated genetic alterations leading to progression. Image includes the points at which specific genetic mutations and alteration occur in the adenoma-carcinoma pathway.**



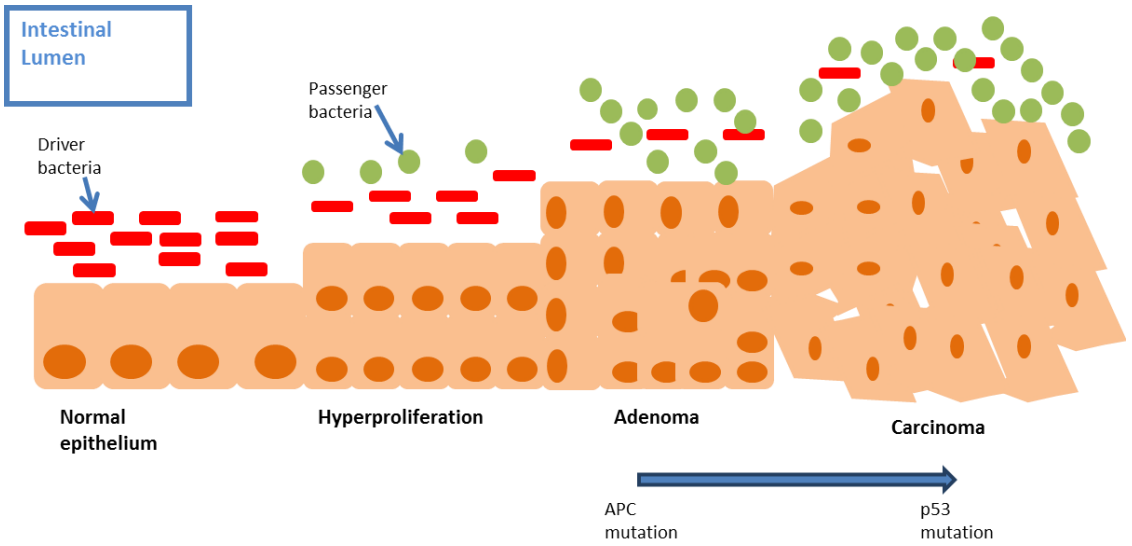

As previously described, the adenoma-carcinoma pathway is pivotal in the development and progression of colorectal cancer (Figure 38). The key genetic abnormalities which occur during this process have been well-described, with some being used to predict outcome and response to treatment. What is less-well understood is the role that the gut microbiota may play within the adenoma-carcinoma pathway. Microbes can potentially have an oncogenic impact in a number of different ways, from DNA integration, affecting genomic instability and resistance to cell death, enhancing proliferative signalling and receptor engagement resulting in pathway signalling[347]. Toxins produced by certain bacteria residing in the large bowel that act as defence strategies for the individual bacteria can cause DNA damage, either as a direct consequence of the toxin or indirectly through the formation of oxygen and nitrogen reactive species (ROS and NOS), examples of such toxins as *Bacteroides fragilis* toxin (Bft) produced *by B. fragilis* and cytolethal distending toxin (CDT)

produced by several Proteobacteria, including *Campylobacter* spp[348,349]. CDT is an example of a toxin that causes direct DNA damage. Bft acts indirectly by eliciting high levels of ROS. Chronically high levels of ROS can outpace a hosts DNA repair ability leading to DNA damage and mutations[347].

Along with the effects of toxins, several microbes can engage directly with host cell pathways, which in turn result in carcinogenesis. *Fusobacterium nucleatum* is able to evoke such a response. It expresses FadA, a bacterial cell surface adhesion component that binds to host E-cadherin, leading to β-catenin activation and cell proliferation. This further emphasises the point that a loss of appropriate boundaries and barrier maintenance between host and microbe is a critical step in the development of some tumours[347,350]. *F. nucleatum* has also been implicated in the chronic inflammation pathways, particularly through its activation of NF-Kappa B. Once the gut barrier is breached, innate and adaptive immune responses are activated, resulting in an inflammatory response. Numerous cytokines and chemokines are released, including IL-23, IL-6, IL-8 and TNF-alpha. Subsequent pathway activations contribute to tumour growth and progression. Evidence suggests that there is preferential selection of bacterial which can thrive in an environment of inflammation and invasion[347,351].

Deep-sequencing technology has allowed further exploration of the role that the gut microbiota plays in the development and progression of colorectal cancer and qualitative and quantitative analysis of microbes found on tumour and healthy tissue[299,352,353]. One of the most striking features of such work is the enrichment of *Fusobacterium* spp. in the on-tumour samples. Histological analysis shows that the bacteria can invade the cancer cells and actually be transported with the cell during metastatic spread. It is clear that such bacteria play a key role but the question of when and how has led to the development of the bacterial driver-passenger model of colorectal cancer (Figure 39)[343].

**Figure 39: A bacterial driver-passenger model of colorectal cancer**




First, certain indigenous intestinal bacteria drive the epithelial DNA damage that contributes to the initiation of colorectal cancer (termed bacterial drivers). Second, tumour development induces intestinal niche alterations that favour the proliferation of opportunistic bacteria (termed bacterial passengers). It should be noted that this driver–passenger model implies that, in contrast to driver mutations in the genomes of cancerous cells, bacterial drivers may disappear from cancerous tissue as they are outcompeted by passenger bacteria with a growth advantage in the tumour microenvironment[343]. Therefore sequencing analysis of on-tumour bacteria appears to be dominated by the passenger bacteria.

### 4.5.1 Bacterial Drivers

A number of bacteria have been suggested as candidate bacterial drivers of colorectal cancer. Many of these bacteria have the production of DNA-damaging compounds in common. Organisms with such features include *Enterococcus faecalis*, which produces extracellular superoxide; this is converted to hydrogen peroxide, leading to DNA damage in colonocytes [354,355]. Certain strains of *Escherichia coli* harbour the polyketide synthetase island which encodes a genotoxin called colibactin. This toxin can induce single-strand DNA breaks, subsequent activation of DNA damage induced pathways increases the mutation rate in infected cells[356]. Others include *B. fragilis* and its production of *Bacteroides fragilis* toxin (Bft) and *Campylobacter showae* related synthesis of cytolethal distending toxin.

The direct action of such bacteria and their toxins is augmented by the induction of chronic inflammation, predominately driven by IL-17 from T-helper cells in the lamina propria. Subsequent

activation of STAT3 and IL-6 appears to have a central role in the development and progression of colorectal cancer, via their pro-proliferative, anti-apoptotic and pro-angiogenic properties[343,357]. Increased levels of such pro-inflammatory mediators, including IL-6 and IL-8 have been reported in studies examining the role of *Campylobacter* spp in the carcinogenesis of colorectal cancer[358]. Such species of bacteria are thought to be relatively rare in healthy human colonic tissue, but have been shown to be over-expressed in non-cancerous mucosa of those with colorectal cancer. *Campylobacter* spp have also been reported to be over-represented in patients with inflammatory bowel disease[349]. Therefore, persistent low-grade colonization with such organisms could increase an individual's susceptibility to colorectal cancer by inducing an asymptomatic but chronic inflammatory response in the colonic mucosa.

**4.5.2 Candidate bacterial passengers**

Bacterial passengers of colorectal cancer are defined as gut bacteria that are relatively poor colonizers of a healthy intestinal tract but have a competitive advantage in the tumour microenvironment, allowing them to outcompete bacterial drivers of CRC[343]. A series of studies have highlighted *Fusobacterium nucleatum* as the most common passenger bacterium, being significantly over-represented on cancer tissue[299,300,353]. This over-representation does not however prove causality. Rubinstein *et al* demonstrated the importance of the adhesion molecule FadA for the stimulation of colorectal cancer. FadA binds vascular endothelial-cadherin on endothelial cells, causing increased endothelial cell permeability thus allowing bacteria to penetrate, a likely mechanism used by *F. nucleatum* for systemic dissemination. This demonstrates that *F. nucleatum* binds to, and invades, both normal and cancerous epithelial cells via FadA binding to epithelial (E)-cadherin. This binding leads to growth stimulation of human colorectal cancer cells but not the non-cancerous cells. FadA binding to E-cadherin on colorectal cancer cells activates β-catenin-regulated transcription, resulting in increased expression of proto-oncogenes cyclin D1 and c-Myc, Wnt and inflammatory genes NF-kappa B, IL-6, IL-8, and IL-18[350,359]. Kostic *et al* showed that *F. nucleatum* was able to induce tumour cell proliferation and selectively recruit tumour-infiltrating myeloid cells to promote tumourigenesis in APC+/- mice. Supporting its role further along the adenoma-carcinoma pathway, rather than initiating the process like the driver bacteria[352]. *F. nucleatum* does not induce colitis or enteritis, therefore it is not able to play a role in inflammation-associated intestinal carcinogenesis, which supports its position as a passenger bacteria in the model.

The ongoing development and progression of colorectal cancer is therefore beneficial to *F. nucleatum* as it out-competes other colonic microbes in such a setting. It is an efficient producer of butyrate via the fermentation of fibre. Much evidence exists to support the beneficial impact of

SCFAs in the colon, acting as an energy source for colonocytes and having an anti-inflammatory effect. However, in the context of this driver-passenger model butyrate may convey a negative impact. Belcheva *et al*, using a murine model with *Msh* and *APC* mutations, demonstrated enhanced tumourigenesis in the presence of butyrate. This may be one way in which *F. nucleatum* maintains its preferential environment[360]. During our study, we found a significant production of butanoic acid in those samples containing pure *F. nucleatum* and 24 hour co-culture. The presence of butyrate during chronic inflammation has been shown to induce apoptosis and, therefore, to reduce tumour development, suggesting that *F. nucleatum* has no in the initiation of the adenoma-carcinoma pathways. However, it may promote its continuation once the pathway is initiated and in doing preferentially selects its own existence.Other compounds that were found to be significantly more abundant in the *F. nucleatum* dominated samples were phenol and two sulphides. Phenol has been demonstrated to promote carcinogenesis in the colon and has been shown to be most abundant in the distal colon where tumour development in most common[287]. Sulphides have been shown to be endogenously produced by colorectal cancer cell via cystathionine b-synthase (CBS). Hydrogen sulphides follow a biphasic dose-response: The effects of some sulphides range from physiological, cytoprotective effects, which occur at low concentrations, to cytotoxic effects, which are typically apparent only at higher concentrations[361]. Sulphides have been demonstrated to promote angiogenesis, the proliferation of endothelial cells, fibroblasts, hepatocytes, and various cancer cells[362]. Thus their presence would support the tumour microenvironment and thus the over-population of *F. nucleatum*. Inhibition of sulphide production by colorectal cancer cells has been shown to reduce angiogenesis and overall tumour growth, potentially leading to new therapeutic targets[361].

It is unclear from this study why the representation of *F. nucleatum* and *C. showae* changes over the time of their co-culture. Given the net utilising properties of *C. showae* it may be that it simply uses up energy stores etc and therefore cannot survive. The human gut would, potentially, act as a constant stream of nutrition and energy for colonising bacteria, an environment that is not mimicked by such culturing. Furthermore, if the mix of metabolites produced by the *F. nucleatum* is toxic it may lead to the death of *C. showae*. This hypothesis may explain how *F. nucleatum* is able to preferentially maintain the tumour microenvironment and why it is over-represented on neoplastic tissue.

**4.6 Conclusion**

Co-culture of *F. nucleatum* and *C. showae* alters the production of VOCs when compared to isolated culture of the bacteria. When the co-culture is for 24 hours there is a large shift towards representation by *F. nucleatum*, this is mirrored by the identification of VOCs. *F. nucleatum* has been suggested as a passenger bacterium, as part of the driver-passenger model of colorectal carcinogenesis. Many of the VOCs identified during this study have been shown to preferentially support the tumour microenvironment and thus the ongoing over-representation of *F. nucleatum*.

# Chapter 5

# An investigation of volatile organic compounds emitted from faeces as a biomarker for colorectal neoplasia

## 5.1 Introduction

Colorectal cancer is a major cause of mortality and morbidity in Europe and North America, with an estimated European incidence of 43.5 per 100,000 and a mortality of 19.5 per 100,000 in 2012[1]. The incidence of CRC has increased by 6% in the last decade and it is now the third most common malignancy in the UK. It leads to approximately 15,000 deaths per annum in the UK, with a 50-55% 5 year mortality rate[2]. Colorectal cancer is the subject of national screening programmes in many countries; these aim to detect cancer in its early stages, including pre-malignant adenomatous polyps. There is clear evidence that such programmes reduce the risk of death from colorectal cancer through detection of tumours at an earlier, a more treatable stage and through removal of these precancerous adenomas[363].

### 5.1.1 Aim

The aim of this study was to determine if the headspace volatile organic compounds (VOCs) detected in faecal samples from patients with neoplasia and no neoplasia in the colon are different from each other and whether faecal VOC profilemay be used as a biomarker for the diagnosis and/or monitoring of colorectal neoplasia.

### 5.1.2 Method

SPME headspace extraction, followed by GCMS, was used to assess VOCs emitted from 450mg aliquots of faeces of symptomatic and Bowel Cancer Screening Programme (BCSP) patients undergoing colonoscopy. Patients were assigned to one of three groups following colonoscopy: adenoma, cancer or non-neoplastic. A full description of the methods can be found in Chapter 2.

## 5.2 Results

### 5.2.1 Patient demographics

There were a total of 133 patients who were recruited to the study. The average age was 64.3 years, 56% of patients were male. The mean age increased in patients with no neoplasia, adenoma and

cancer respectively, *p=0.02*. None of the participants reported being a smoker or a vegetarian. Self-reported ethnicity was used, all but one patient was said to be white British. Participation in the BCSP provided the largest proportion of referrals for colonoscopy (31.5%). Full demographics and clinical characteristics can be seen in Table 20: 11 patients were from Sheffield, all of whom had a confirmed diagnosis of colorectal cancer but no other clinical information was available for these patients.

**Table 22: Summary of known demographic data**

| | **Total** | **Non-neoplastic** | **Adenoma** | **Cancer** |
|---|---|---|---|---|
| **Number** | 137 | 60 | 56 | 21 |
| **Mean age (range)** | 64.3 (22-85) | 61.9 (22-85) | 65.6 (41-84) | 72.7 (64-78) |
| **Gender** | | | | |
| Male | 69 | 25 | 36 | 7 |
| Female | 57 | 34 | 20 | 3 |
| **Smoker (Yes)** | 0 | 0 | 0 | 0 |
| **Indication for colonoscopy** | | | | |
| BCSP | 38 | 13 | 22 | 3 |
| IDA | 27 | 16 | 6 | 5 |
| Change in bowel habit-diarrhoea | 16 | 11 | 4 | 1 |
| Surveillance previous neoplasia/FH | 35 | 10 | 24 | 1 |
| IBD assessment/surveillance | 9 | 9 | 0 | 0 |
| GI bleeding | 1 | 1 | 0 | 0 |
| Unknown | 11 | 0 | 0 | 11 |

Table includes diagnosis and indications for original colonoscopy, of patients recruited in Liverpool, Plymouth and Sheffield. N.B. 11 of the cancer samples had no demographic information available.

A third cohort of patients was supplied from Plymouth: all those were undergoing colonoscopy as part of their investigations for iron deficiency anaemia. Samples were produced, handled and stored in line with the methodology employed in Liverpool. This cohort consisted of 12 adenomatous polyp patients and 4 with adenocarcinoma. These samples were not used to build the compound library

and were used as part of the validation process. The mean age of these samples was 73 years. Only when specified in this chapter were these samples utilised in the analysis.

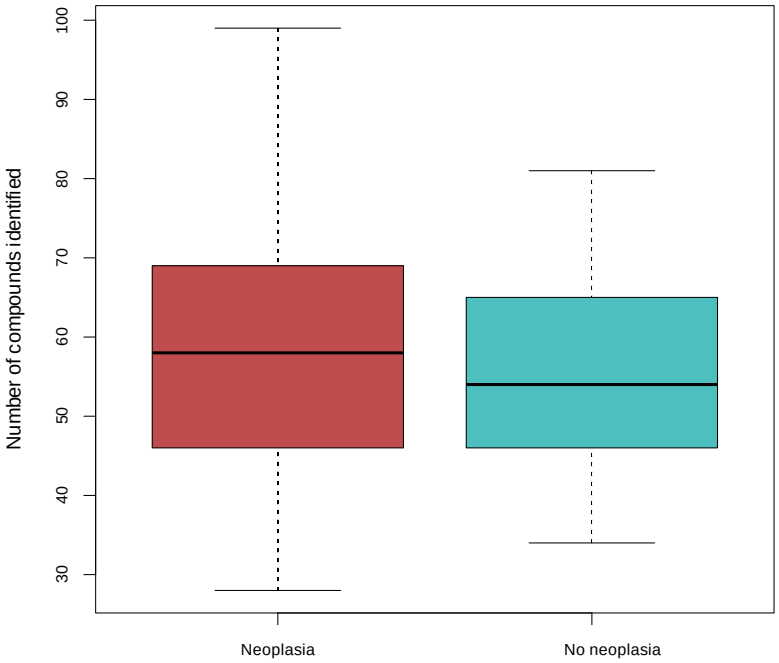
### 5.2.2 VOC characterisation

A total of 162 VOCs were identified across all samples. The mean number of VOCs identified in the entire cohort was 56.7, with no significant difference in those with or without neoplasia, *p=0.2*, (Table 23, Figure 40).

**Table 23: Mean number of VOCs identified in patients with and without colonic neoplasia.**

| | No- neoplasia | Neoplasia |
|---|---|---|
| **Number of patients** | 60 | 73 |
| **Mean number of VOCs (SD)** | 58.1 (15.6) | 55.2 (12.1) |
| **Coefficient of variation (CV)** | 0.2 | 0.2 |

**Figure 40: Box and whisker plot representing the number of VOCs identified in those with and without colonic neoplasia.**



**5.3 Univariate analysis in order to identify potential biomarker candidate/s**

Three technical processes were employed in order to identify and compare VOCs in order to identify compounds that could be taken forward for biomarker assessment. A combination of Metab/R[320], Metaboanalyst[324] and XCMS were used to assess fold change, prevalence and relative abundance of VOCs emitted from the samples. When Metaboanalyst was used the data was log transformed and normalised by median. These methods allow for the application of ANOVA, Student's t test and Fisher's exact test. A number of different comparisons were subsequently performed, these were:

- No neoplasia (control) versus neoplasia
- No neoplasia versus neoplasia in the form of adenomatous polyps
- No neoplasia versus neoplasia in the form of colonic adenocarcinoma
- No neoplasia versus a single adenomatous polyps > 1cm is size
- No neoplasia versus patients with > 4 individual adenomatous polyps of any size
- No neoplasia versus neoplasia in the form of adenomatous polyps

The use of multiple independent analytical systems ensured robust identification of the potential VOC biomarkers. Q values were generated to account for false discovery. Results from ANOVA of no neoplasia, adenoma and cancer can be seen in Table 24.

**Table 24: VOCs potentially associated with no neoplasia, adenoma or cancer.**

| VOC (inc RT) | p value | q value |
|---|---|---|
| 27.19_Pentane, 2,3,4-trimethyl- | 0.001 | 0.09 |
| 32.25_menthol | 0.001 | 0.09 |
| 19.60_Butanoic acid, 3-methyl-, ethyl ester | 0.001 | 0.09 |
| 20.53_p-Xylene | 0.005 | 0.18 |
| 28.53_Benzeneacetaldehyde | 0.006 | 0.18 |
| 25.66_Propanoic acid, 2-methyl-, 2-methylbutyl ester | 0.006 | 0.18 |
| 14.07_Propanoic acid, anhydride | 0.01 | 0.38 |
| 18.18_Acetic acid, butyl ester | 0.02 | 0.51 |
| 27.52_Butanoic acid, 4-pentenyl ester | 0.03 | 0.51 |
| 7.59_Isopropyl alcohol | 0.03 | 0.51 |
| 22.11_Bicyclo[3.1.0]hex-2-ene, 4-methyl-1-(1-methylethyl)- | 0.03 | 0.51 |
| 24.45_2,6-Octadiene, 2,6-dimethyl- | 0.03 | 0.51 |
| 13.28_1-Butanol | 0.04 | 0.51 |
| 22.06_alpha-Phellandrene | 0.04 | 0.51 |

ANOVA was used to compare no neoplasia, adenoma and cancer. q values representative of false discovery rate. Analysis was performed using Metaboanalyst. RT= Retention time.

A number of the compounds found in Table 3 were seen to be significantly different (p=<0.05) across the other comparisons and modes of analysis. In particular, these were benzeneacetaldehyde, 2,3,4-trimethyl-pentane, menthol, isopropyl alcohol, butanoic acid based compounds, esters of acetic acid and sulphides. However, none remained significant after correction for multiple comparisons.

Those VOCs identified to be significant different when comparing no neoplasia to adenoma and cancer can be seen in Tables 25 and 26. Again, most were not significant after correction for multiple comparisons.

**Table 25: VOCs potentially associated with no neoplasia or adenoma.**

| VOC (RT) | p value | q value |
|---|---|---|
| 16.69_Methyl isovalerate | 0.01 | 0.73 |
| 27.52_Butanoic acid, 4-pentenyl ester | 0.02 | 0.73 |
| 30.30_Pentanoic acid, 3-methylbutyl ester | 0.02 | 0.73 |
| 27.19_Pentane, 2,3,4-trimethyl- | 0.03 | 0.73 |
| 28.53_Benzeneacetaldehyde | 0.03 | 0.73 |
| 30.78_p-Cresol | 0.03 | 0.73 |
| 22.06_alpha-Phellandrene | 0.03 | 0.73 |

Table generated from Student's t test comparing no neoplasia and adenoma. q value representative of false discovery rate. Analysis performed using Metaboanalyst. RT= Retention time

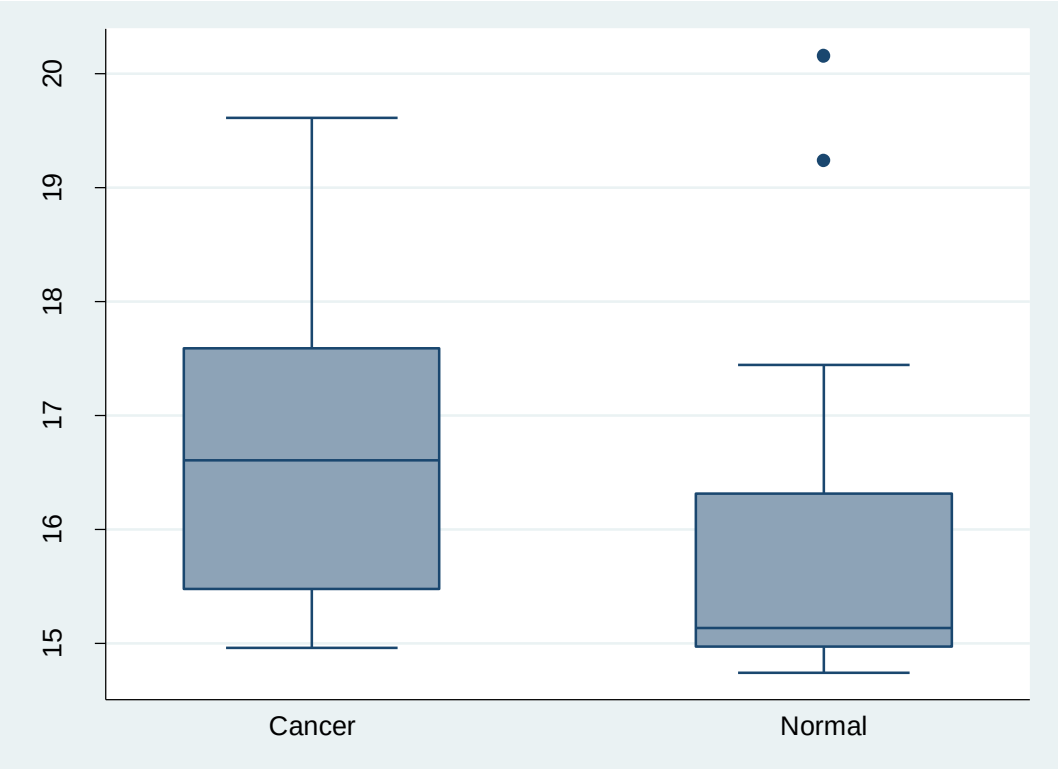
**Table 26: VOCs with p value <0.05 when comparing no neoplasia and cancer. False discovery rate applied to correct for multiple comparisons, represented as q value.**

| VOC (RT) | p value | q value |
|---|---|---|
| 7.59_Isopropyl alcohol | <0.0001 | 0.004 |
| 17.98_2-Hexanone | 0.01 | 0.77 |
| 19.60_Butanoic acid, 3-methyl-, ethyl ester | 0.03 | 0.77 |
| 19.07_Butanoic acid, 1-methylethyl ester | 0.03 | 0.77 |
| 22.77_Pentanoic acid, 1-methylethyl ester | 0.03 | 0.77 |
| 20.53_p-Xylene | 0.03 | 0.77 |
| 15.67_Propanoic acid, 1-methylethyl ester | 0.04 | 0.77 |
| 32.25_menthol | 0.05 | 0.77 |

Table generated from Student's t test comparing no neoplasia and cancer. q value representative of false discovery rate. Analysis performed using Metaboanalyst. RT= Retention time.

After correction for false discovery rate isopropyl alcohol shows the most promise for the detection of neoplasia, particularly adenocarcinoma.

**Figure 41: Box and whisker plot for isopropyl alcohol concentration found in faecal samples derived from patient with adenocarcinoma of the colon and no colonic neoplasia.**

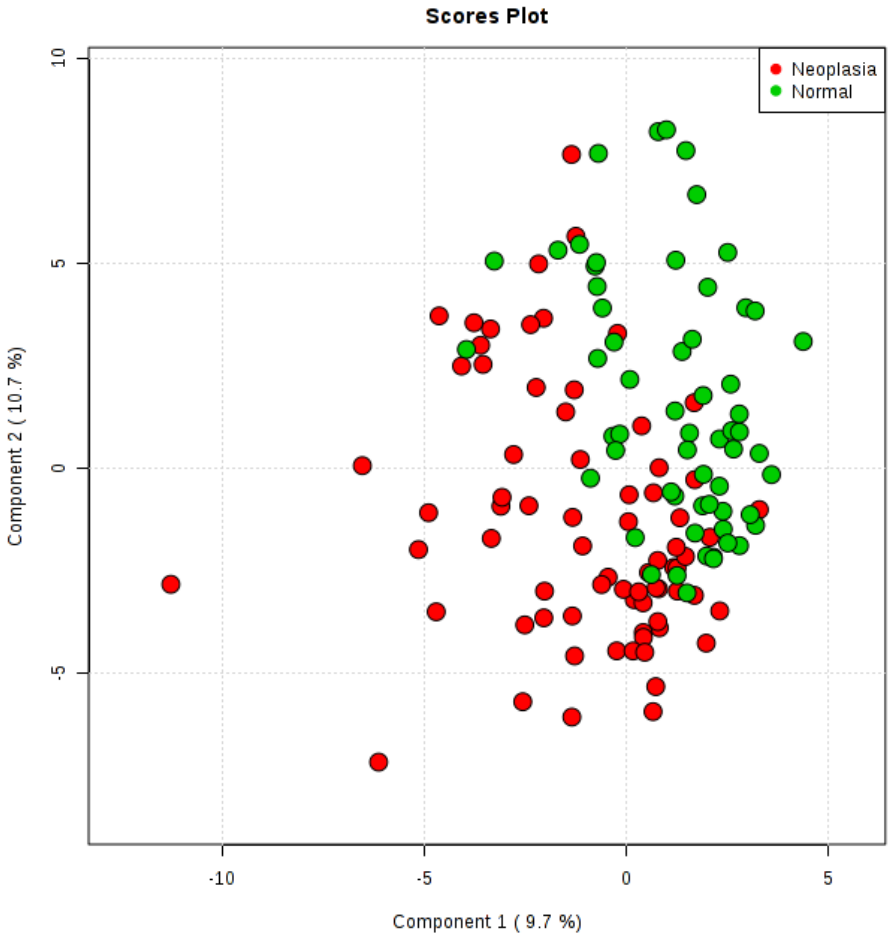



Normal represents those without colonic neoplasia. The adenocarcinoma cohort includes those cancer samples from Plymouth.

### 5.4 Analysis of potential biomarkers for colorectal cancer identified from univariate analysis

Initially, PLS-DA was employed to assess the ability to distinguish between the presence and absence of colonic neoplasia. This failed to show a significant degree of separation (Figure 42). ROC analysis did not achieve an AUROC curve great enough to be considered adequate for a diagnostic tool. Benzeneacetaldehyde achieved the greatest AUROC, 0.67. Analysis was then focused upon higher risk neoplastic disease, namely adenocarcinoma, a single adenomatous polyp > 1cm in size and individuals with > 4 adenomatous polyps of any size.

**Figure 42: Partial least squared discriminant analysis for all identified VOCs found in those with and without colonic neoplasia.**

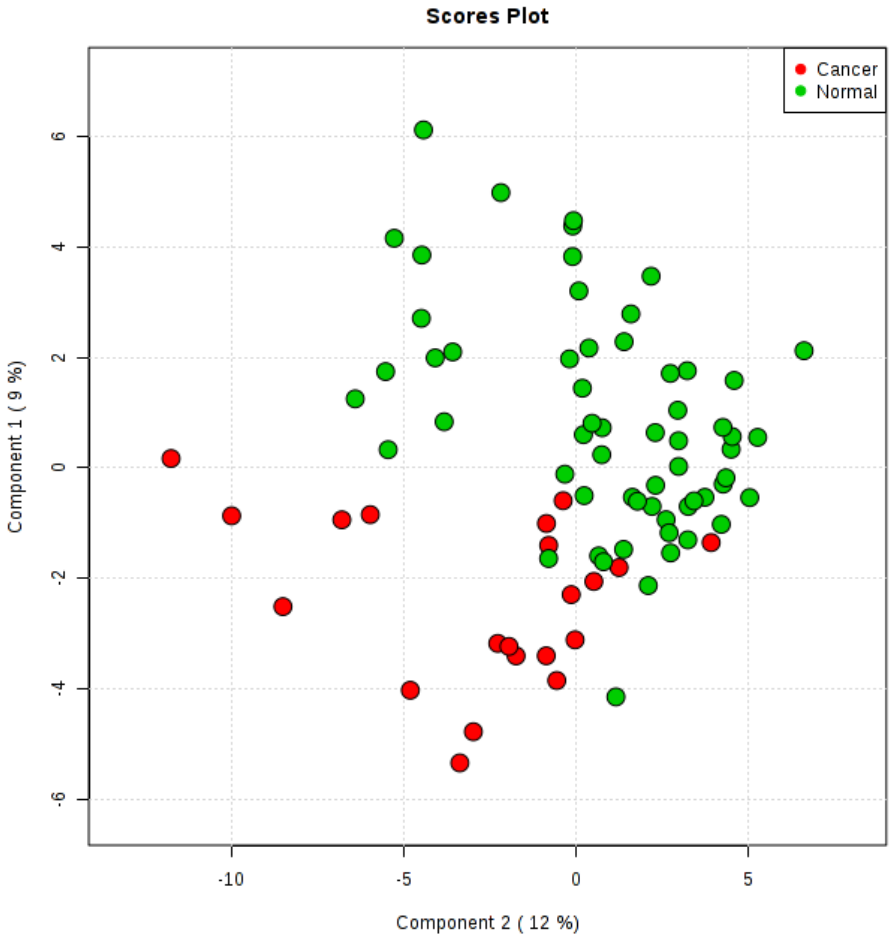



Normal refers to those without neoplasia. Analysis based upon all VOCs found.

### 5.4.1 VOCs as a biomarker for colonic adenocarcinoma

Isopropyl alcohol was further investigated because of the q value found earlier. PLS-DA comparing those without neoplasia and adenocarcinoma (cancer) showed a degree of separation that suggested some diagnostic utility (Figure 43). ROC analysis illustrated the potential to diagnosis adenocarcinoma of the colon using isopropyl alcohol, with an AUROC of 0.76 (Table 27). When considering isopropyl alcohol in isolation, assessment of sensitivity and specificity can be made according to differing cut-offs. The optimal cut-off gave a sensitivity of 83% and specificity of 71% (Figure 44).

**Figure 43: PLS-DA comparing those with adenocarcinoma of the colon and no colonic neoplasia.**



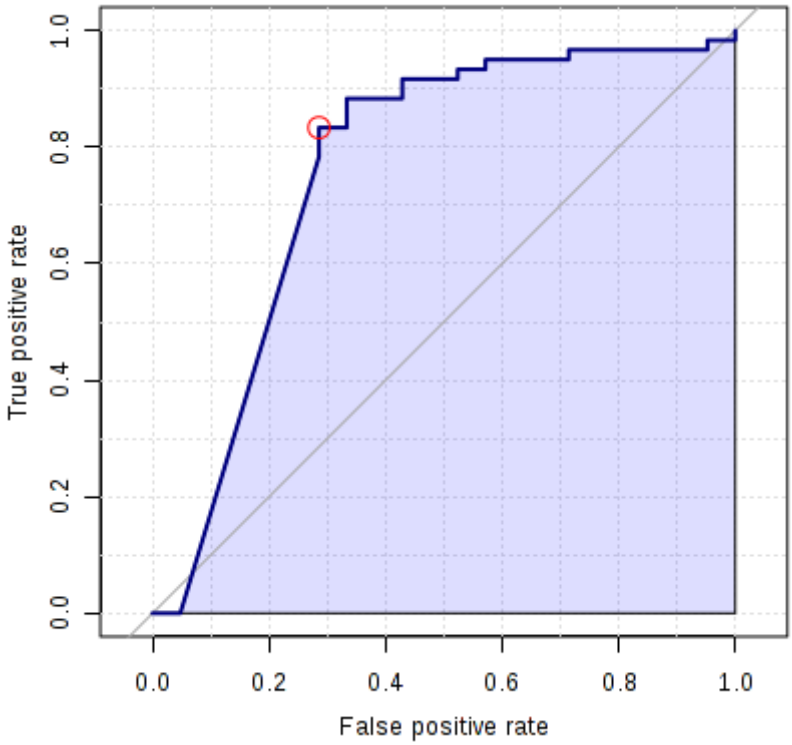

Normal refers to those without neoplasia.  Analysis based upon all VOCs found.

**Table 27: AUROC results for the VOCs emitted when comparing those with adenocarcinoma of the colon and no colonic neoplasia.**

| VOC (RT) | AUROC | p value |
|---|---|---|
| 7.59_Isopropyl alcohol | 0.76 | <0.0001 |
| 17.98_2-Hexanone | 0.66 | 0.01 |
| 20.53_p-Xylene | 0.66 | 0.03 |
| 22.77_Pentanoic acid, 1-methylethyl ester | 0.65 | 0.03 |
| 19.07_Butanoic acid, 1-methylethyl ester | 0.65 | 0.03 |
| 29.04_2-Nonanone | 0.64 | 0.05 |
| 19.60_Butanoic acid, 3-methyl-, ethyl ester | 0.63 | 0.02 |
| 32.25_menthol | 0.62 | 0.04 |

Those with an AUROC of >0.6 have been included in the table.

**Figure 44: ROC curve for isopropyl alcohol when comparing those with adenocarcinoma of the colon and no colonic neoplasia.**




The red circle represents optimal cut-off generating sensitivity of 83% and specificity of 71%.

Calculating ratios of all possible metabolite pairs and then choosing top ranked ratios, based on p values, allowed for further biomarker assessment. Only those achieving an AUROC greater than isopropyl alcohol alone are reported (Table 28). All 7 combinations reported contain isopropyl alcohol.

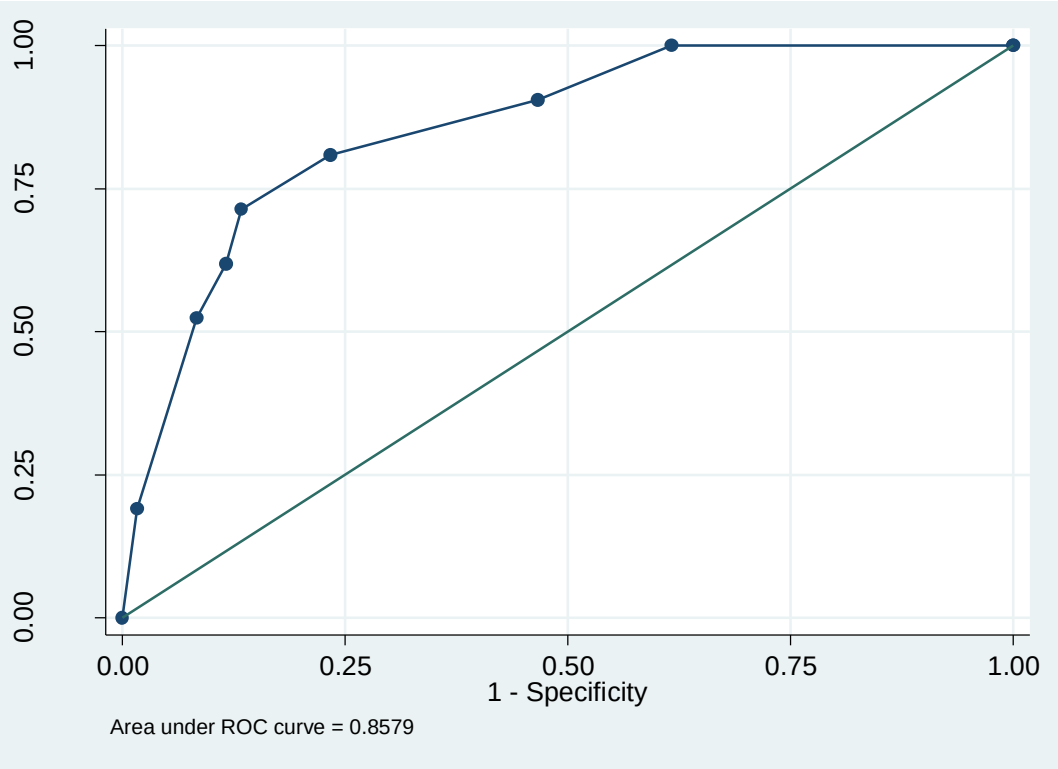
**Table 28: AUROC results for the VOCs emitted when using a comparison of ratios for those with adenocarcinoma of the colon and no colonic neoplasia.**

| VOC ratio combination | AUROC | 95% CI | Optimal Sensitivity % | Optimal Specificity % |
|---|---|---|---|---|
| 21.35_Butanoic acid, 3-methyl-/7.59_Isopropyl alcohol | 0.82 | 0.71-0.92 | 81 | 76 |
| 32.25_menthol/7.59_Isopropyl alcohol | 0.82 | 0.7-0.91 | 85 | 71 |
| 29.04_2-Nonanone/7.59_Isopropyl alcohol | 0.81 | 0.7-0.91 | 88 | 57 |
| 18.20_Propanoic acid, 2-methyl-/7.59_Isopropyl alcohol | 0.81 | 0.7-0.91 | 80 | 76 |
| 24.13_Decane/7.59_Isopropyl alcohol | 0.81 | 0.68-0.9 | 81 | 71 |
| 22.01_Acetic acid, pentyl ester/7.59_Isopropyl alcohol | 0.8 | 0.69-0.89 | 83 | 71 |
| 28.53_Benzeneacetaldehyde/7.59_Isopropyl alcohol | 0.79 | 0.65-0.91 | 86 | 76 |

### 5.4.2 VOCs as a biomarker for colonic adenocarcinoma- Qualitative analysis

Logistic regression analysis for compound presence, identified 3 VOCs as potential biomarkers. These were isopropyl alcohol, 2-hexanone and butanoic acid, 3-methyl-,ethyl ester. Using all 3 VOCs as a biomarker panel achieves an AUROC of 0.86 (Figure 45). Using this model, there is a 6-fold increased chance of colorectal cancer if all 3 VOCs are present in a patient's stool sample.

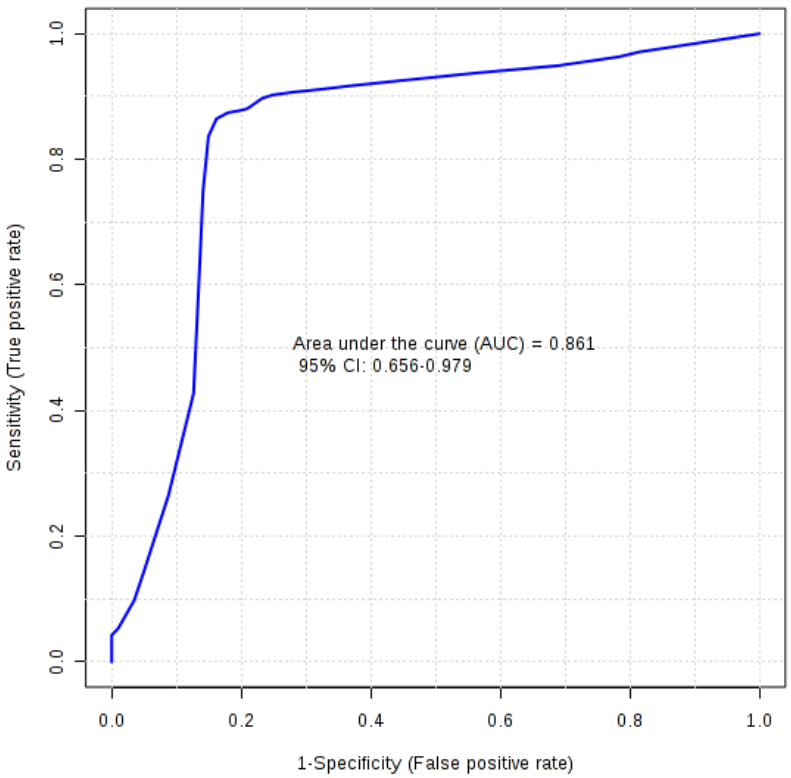
**Figure 45: ROC curve using the qualitative biomarker panel of isopropyl alcohol, 2-hexanone and butanoic acid, 3-methyl-,ethyl ester for the diagnosis of colorectal cancer.**




### 5.4.2.1 Validation of quantitative biomarker model for colorectal cancer

A hold-out technique was applied to the 81 samples (21 cancer and 60 controls) in order to validate the combination of butanoic acid, 3-methyl-/isopropyl alcohol as a biomarker for colorectal cancer. Fifty percent of each cohort were held back, the remaining samples were then subjected to testing with the butanoic acid, 3-methyl-/isopropyl alcohol combination via logistic regression and a 10-fold cross validation method (Figure 50). When the model is then applied to those samples held out, the AUROC is 0.82, sensitivity 87.9% (95% CI 0.87-0.99) and specificity 84.6% (95% CI 0.65-1.0).

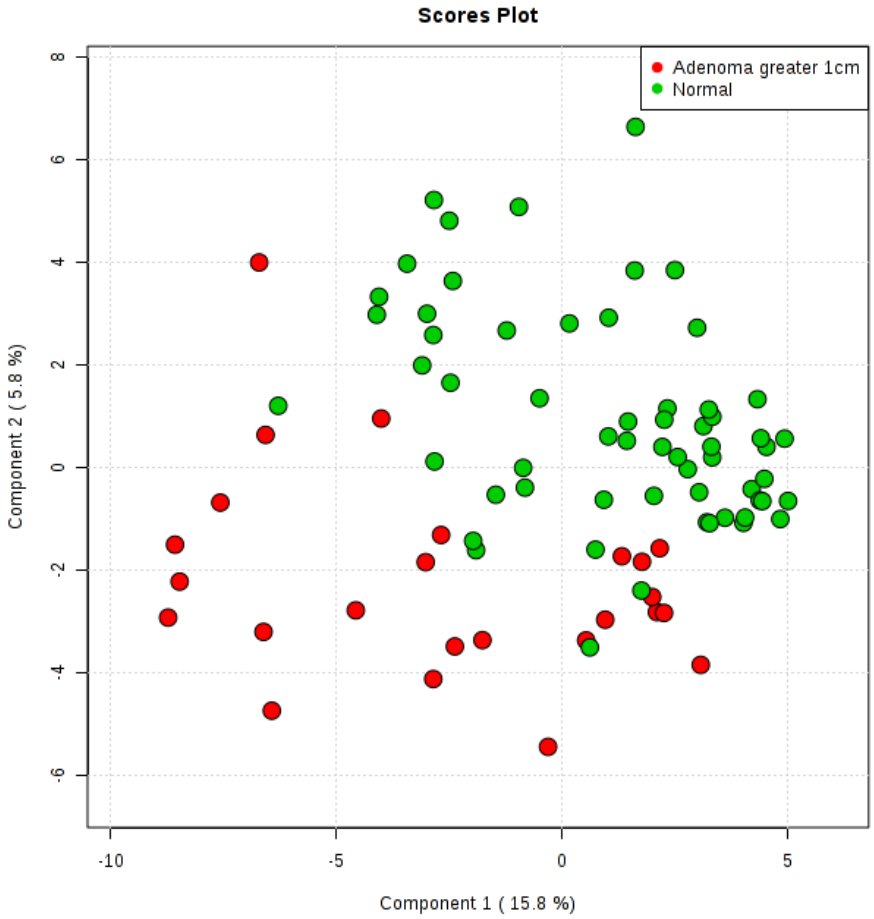
**Figure 46: ROC curve using logistic regression and 10-fold cross-validation, based on the combination of butanoic acid, 3-methyl-/isopropyl alcohol, for the diagnosis of colorectal cancer.**




### 5.4.3 VOCs as a biomarker for adenomatous polyps > 1cm in size - quantitative analysis

PLS-DA showed separation between those without neoplasia and those with an adenomatous polyp > 1cm in size (Figure 47). Subsequent univariate ROC analysis demonstrated an AUROC of 0.7 for 2,3,4-trimethylpentane, with a potential sensitivity of 78% and specificity of 68% (Table 29, Figure 48). These compounds were then taken forward for further biomaker assessment using the combined ratio approach (Table 30).

**Figure 47: PLS-DA comparing those with no colonic neoplasia and a single adenomatous polyp > 1cm in size**



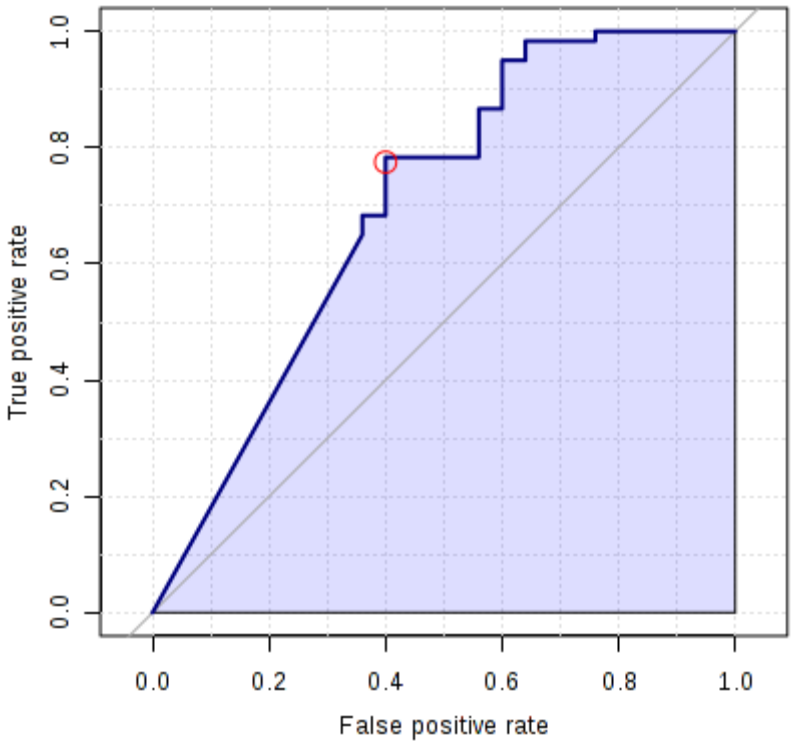

Based on all VOCs present and generated via Metaboanalyst.

**Table 29: AUROC result for the VOCs emitted when comparing no colonic neoplasia and a single adenomatous polyp > 1cm in size.**

| VOC (RT) | AUROC | p value |
|---|---|---|
| 27.19_Pentane, 2,3,4-trimethyl- | 0.70 | <0.0001 |
| 27.52_Butanoic acid, 4-pentenyl ester | 0.67 | <0.0001 |
| 28.53_Benzeneacetaldehyde | 0.70 | 0.02 |
| 27.06_Butanoic acid, 3-methylbutyl ester | 0.67 | 0.02 |
| 32.25_menthol | 0.66 | 0.01 |
| 35.31_2-Undecanone | 0.66 | 0.01 |
| 26.96_Pentanoic acid, 2-methylpropyl ester | 0.66 | 0.01 |
| 29.39_Nonanal | 0.65 | 0.03 |
| 28.33_Pentanoic acid, butyl ester | 0.65 | 0.01 |

Those with an AUROC of >0.6 have been included in the table.

**Figure 48: ROC curve for Pentane, 2,3,4-trimethyl- following univariate analysis, when comparing those with no colonic neoplasia and a single adenomatous polyp > 1cm in size.**




The red circles represent the optimal cut off point giving sensitivity of 78% and specificity of 68%.

Following the comparison of abundance ratios, the greatest AUROC achieved was 0.78, this was the combination of benzeneacetaldehyde and 2-undecanone, demonstrating a sensitivity of 76% and specificity of 72% (Table 30). When combined with a number of other VOCs 2,3,4-trimethylpentane was noted to have an improved AUROC.

**Table 30: AUROC result for the VOCs emitted when using a comparison of ratios for those with no colonic neoplasia and those with a single adenomatous polyp >1cm in size.**

| VOC ratio combination | AUROC | 95% CI | Optimal Sensitivity % | Optimal Specificity % |
|---|---|---|---|---|
| 28.53_Benzeneacetaldehyde/35.31_2-Undecanone | 0.78 | 0.68-0.87 | 76 | 72 |
| 27.19_Pentane, 2,3,4-trimethyl-/32.25_menthol | 0.77 | 0.67-0.86 | 80 | 64 |
| 19.60_Butanoic acid, 3-methyl-, ethyl ester/ 27.06_Butanoic acid, 3-methylbutyl ester | 0.76 | 0.65-0.85 | 81 | 56 |
| 29.39_Nonanal/35.31_2-Undecanone | 0.76 | 0.65-0.84 | 76 | 52 |
| 32.25_menthol/35.31_2-Undecanone | 0.75 | 0.63-0.84 | 68 | 72 |
| 19.60_Butanoic acid, 3-methyl-, ethyl ester/ 27.19_Pentane, 2,3,4-trimethyl- | 0.74 | 0.63-0.84 | 68 | 60 |

### 5.4.4 VOCs as a biomarker for > 4 individual adenomatous colonic polyps- quantitative analysis

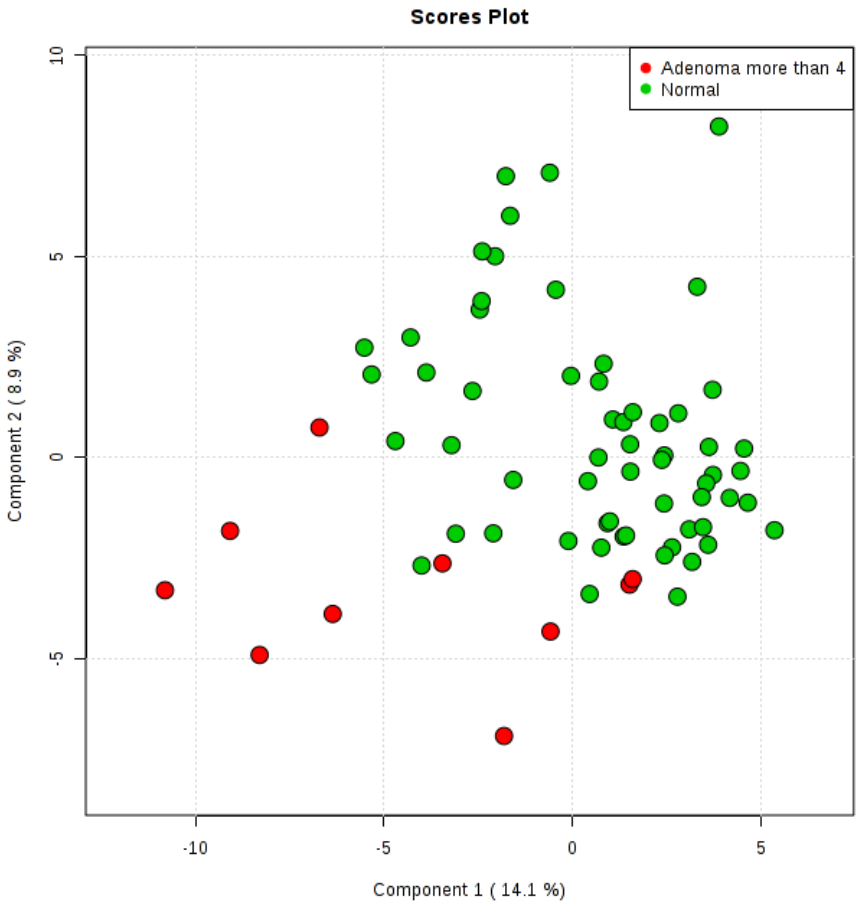
Univariate analysis was again performed, including correction for false discovery rate, in order to identify potential candidate biomarkers (p=<0.05). Five of the nine VOCs identified as potential candidates in the analysis of polyps >1cm were again identified in this sub-analysis. These were pentane, 2,3,4-trimethyl-, benzeneacetaldehyde, butanoic acid, 4-pentenyl ester, butanoic acid, 3-methylbutyl ester and menthol. After false discovery rate correction had been applied none of these VOCs achieved a q value of <0.05, despite this there did appear to be a good degree of separation on the PLS-DA (Figure 53). Hexanoic acid, ethyl ester demonstrated the most potential, as it had the lowest p and q values (0.001 and 0.09, respectively) (Table 31).

**Table 31: VOCs with p value <0.05 when comparing no neoplasia and those with > 4 individual polyps of any size.**

| VOC (RT) | p value | FDR |
|---|---|---|
| 25.03_Hexanoic acid, ethyl ester | 0.001 | 0.09 |
| 27.19_Pentane, 2,3,4-trimethyl- | 0.004 | 0.26 |
| 31.50_Hexanoic acid, butyl ester | 0.01 | 0.26 |
| 28.53_Benzeneacetaldehyde | 0.01 | 0.26 |
| 28.72_Propanoic acid, hexyl ester | 0.01 | 0.26 |
| 27.52_Butanoic acid, 4-pentenyl ester | 0.01 | 0.26 |
| 25.61_Acetic acid, hexyl ester | 0.01 | 0.26 |
| 25.49_Benzaldehyde | 0.02 | 0.26 |
| 33.32_Isopentyl hexanoate | 0.02 | 0.26 |
| 18.57_Methyl valerate | 0.02 | 0.26 |
| 27.06_Butanoic acid, 3-methylbutyl ester | 0.02 | 0.26 |
| 23.39_Methional | 0.03 | 0.34 |
| 26.33_Hexanoic acid | 0.03 | 0.34 |
| 30.32_Ethanone, 1-(2-methyl-1-cyclopenten-1-yl)- | 0.03 | 0.34 |
| 30.30_Pentanoic acid, 3-methylbutyl ester | 0.04 | 0.35 |
| 32.25_menthol | 0.04 | 0.35 |
| 12.47_Butanal, 3-methyl- | 0.04 | 0.36 |
| 17.98_2-Hexanone | 0.04 | 0.36 |

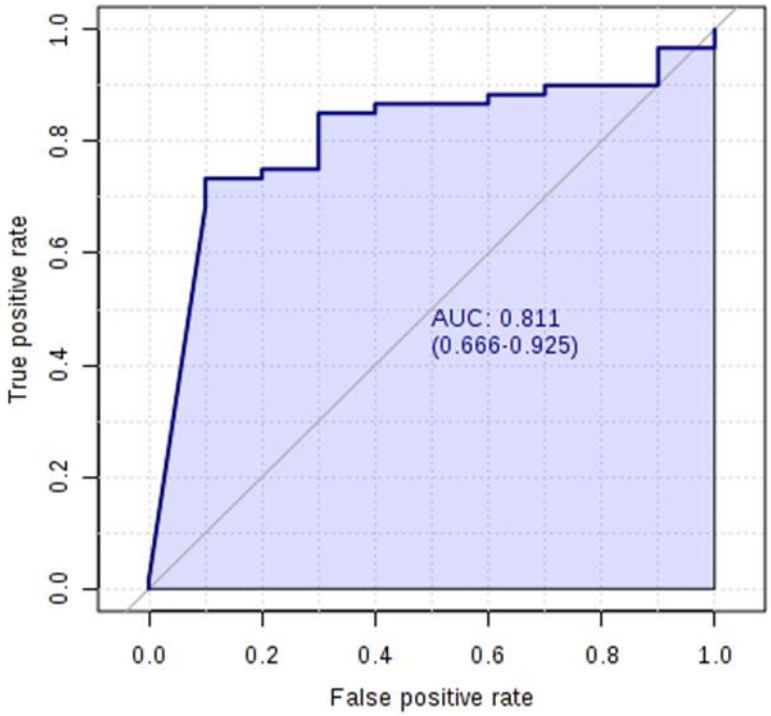
Table generated from Students t test, q value generated for false discovery rate. Analysis performed using Metaboanalyst. RT= Retention time.

**Figure 49: PLS-DA using all the identified VOCs, comparing those with no neoplasia against those with > 4 individual polyps of any size.**




Using univariate ROC analysis, 6 VOCs achieved an AUROC of >0.7, with two being >0.75. These were hexanoic acid ethyl ester (0.81), butanal, 3-methyl- (0.75), pentane, 2,3,4-trimethyl- (0.74), methyl valerate (0.73), hexanoic acid (0.72), benzeneacetaldehyde (0.71). Hexanoic acid ethyl ester had the greatest AUROC, achieving a sensitivity of 85% and specificity of 70% (Figure 50).

**Figure 50: ROC curve for hexanoic acid ethyl ester**




Box and whisker plot demonstrating the greater abundance of hexanoic acid ethyl ester in those with > 4 individual adenomatous polyps.

Paying particular attention to hexanoic acid ethyl ester, again further assessment of biomarker potential was conducted with ratio comparisons. Only those achieving an AUROC of greater than, or equal to, that of hexanoic acid alone are shown (Table 32).

**Table 32: AUROC result for the VOCs emitted when using a comparison of ratios for those with no colonic neoplasia and those with > 4 individual polyps of any size.**

| VOC ratio combination | AUROC | 95% CI | Optimal Sensitivity % | Optimal Specificity % |
|---|---|---|---|---|
| 19.60_Butanoic acid, 3-methyl-, ethyl ester/ 25.03_Hexanoic acid, ethyl ester | 0.85 | 0.74-0.94 | 80 | 80 |
| 25.03_Hexanoic acid, ethyl ester/ 27.95_Cyclohexanecarboxylic acid, methyl ester | 0.85 | 0.72-0.94 | 90 | 80 |
| 25.03_Hexanoic acid, ethyl ester/ 30.32_Ethanone, 1-(2-methyl-1-cyclopenten-1-yl)- | 0.85 | 0.73-0.94 | 78 | 90 |
| 16.69_Methyl isovalerate/25.03_Hexanoic acid, ethyl ester | 0.84 | 0.73-0.93 | 78 | 90 |
| 25.03_Hexanoic acid, ethyl ester/32.25_menthol | 0.84 | 0.73-0.93 | 78 | 70 |
| 12.47_Butanal, 3-methyl-/17.98_2-Hexanone | 0.84 | 0.71-0.94 | 78 | 80 |
| 25.03_Hexanoic acid, ethyl ester/28.53_Benzeneacetaldehyde | 0.81 | 0.67-0.92 | 85 | 60 |

Table also includes 95% CI and sensitivity and specificity for each VOC combination.

**5.5 Discussion**

Correctly identifying patients to undergo colonoscopy, as part of population based screening, is vital in order to maximise pathology capture and to minimise unnecessary examinations. There is a clear link to improved outcomes from colorectal cancer by the identification of earlier stage colorectal cancer and pre-malignant adenomatous colonic polyps[364]. This study has demonstrated the utility of VOCs emitted from faeces to act as a biomarker for colonic neoplasia, in particular, adenocarcinoma and adenomatous polyps of increasing size and number.

Using a variety of methods and substrates, other studies have suggested a utility of VOC analysis for the diagnosis of colorectal cancer. One such study, from 2015, used selected ion flow tube mass spectrometry (SIFT-MS) to analysis VOCs emitted from faeces of FOBt positive patients,. Comparing patients with no neoplasia and high grade neoplasia, ions probably arising from hydrogen sulphide, dimethyl sulphide and dimethyl disulphide were significantly higher in samples from high risk compared to low risk subjects. The authors reported overall specificity of 78% and 72% sensitivity[365]. Two separate studies, from 2014 and 2013, reported the analysis of VOCs found in urine and breath, respectively. The study examining urine used Field Asymmetric Ion Mobility Spectrometer (FAIMS): 133 patients were included; 83 colorectal cancer patients and 50 healthy controls. Sensitivity and specificity for CRC detection with FAIMS were 88% and 60% respectively[312]. A third technology, in the form of thermal-desorber gas chromatography–mass spectrometry, was used to assess VOCs in the study examining breath. Assessing the pattern of 15 compounds showed a sensitivity of 86%, a specificity of 83% and AUROC of 0·85[311]. More recently, using the same technique, this group described the ability of exhaled VOCs to discriminate between colorectal cancer patients before and after curative surgery[366]. Another study from 2014 reported the utility of a pattern recognition–based detection technique, using VOCs found in faeces. This study did not attempt to identify the individual compounds but focused upon differing patterns. It attempted to identify established colorectal cancer and pre-malignant adenomatous lesions. Faecal VOC profiles of patients with colorectal cancer differed significantly from controls (AUROC, 0.92; sensitivity, 0.85; and specificity, 0.87). Patients with advanced adenomas could also be distinguished from controls (AUROC, 0.79; sensitivity, 0.62; and specificity, 0.86).

Unlike in our work, few studies have attempted to identify the individual compounds responsible for generating the differentiation. Most rely on sensor technology that recognises patterns rather than individual compounds. Identification of the VOCs influencing diagnostic separation allows for biological plausibility to be explored, improving stringency. We also examined the potential of VOCs

for the identification of pre-malignant adenomatous polyps, making our study unique. All of the studies described here are feasibility studies and do not include validation sets, potentially limiting their clinical application. We performed validation of our set with cross-validation methodology. In doing so, we can report diagnostic accuracy for colorectal cancer that is very similar to other studies, but has been through a validation process, AUROC 0.82, sensitivity 87.9% (95% CI 0.87-0.99) and specificity 84.6% (95% CI 0.65-1.0). Moreover, a validated and superior diagnostic utility for pre-malignant adenomatous colonic polyps can be seen in our results.

Population-based screening or a point of care test are the most likely clinical application of such VOC analysis. Despite their relatively low patient acceptance rates, faecal based techniques are currently the most commonly employed i.e. FOBt, either gFOBt or FIT. The gFOBt currently used in the UK BCSP has a sensitivity of 36% and a specificity of 94% for the detection of colorectal cancer[94,95]. To date, there are no controlled trials that demonstrate that FIT are superior to gFOBt or to no screening in terms of reducing colorectal cancer-related mortality in average risk persons. However, a recent observational study from Italy demonstrated a reduction in colorectal cancer-related mortality in regions where screening with FIT was adopted compared with regions where screening had not yet been implemented[367,368] The superiority of FIT over gFOBt is now widely recognised and the European Quality Assurance Guideline on Colorectal Cancer Screening published in 2011 recommends FIT in preference to gFOBt[369,370]. Various countries have adopted FIT into their colorectal cancer screening programmes and the BCSP plans to replace gFOBt with FIT[371]. Comparing the result of our study it would appear that VOCs have a greater diagnostic ability than FOBt for the identification of colorectal cancer and pre-malignant adenomatous polyps. It also appears to have a superior performance when compared to faecal tM2-PK[194], discussed in Chapter 7.

Many of the previous attempts to explore VOCs as a biomarker for colorectal neoplasia have used a pattern recognition system rather than the identification of individual VOCs. Throughout the different elements of our analysis, including factor analysis, key VOCs became evident. In particular these were isopropyl alcohol and butanoic acid based compounds.

The International Agency for Research on Cancer (IARC), part of the World Health Organization (WHO), have classified isopropyl alcohol as a Group 1 carcinogen, meaning there is clear evidence that is carcinogenic to humans[372]. It is clear that isopropyl alcohol is more prevalent and abundant in the faeces of patients with colorectal cancer. Endogenous isopropyl alcohol may arise from the metabolism of non-absorbable carbohydrates (Figure 51 and 52). Fermentation of non-absorbable carbohydrates, by colonic bacteria, is a key source of SCFA such as butyrate, acetate and

propanoate. Along with the metabolism of pyruvate these metabolic pathways are a source of acetate and subsequently acetone. Acetone is converted to isopropyl alcohol by alcohol dehydrogenase, making it a by-product of SCFA metabolism. Many studies have shown the protective effect of SCFA against colorectal cancer, with more recent studies demonstrating the potential negative impact of butyrate as the adenoma-carcinoma pathway progresses. One such study, using a murine model with *Msh* and *APC* mutations, demonstrated enhanced tumourigenesis in the presence of butyrate[360]. If the tumour microenviroment preferentially selects butyrate producing organisms in order to promote its ongoing existence, as is suggested by the driver-passenger model[343], then there will be an increase in butyrate based compounds together with the by-product, isopropyl alcohol, which itself may promote on-going tumourigenesis.

**Figure 51: Schematic representation of the products of fermentation of non-absorbable carbohydrates in the human colon. Adapted from associated KEGG pathway.**

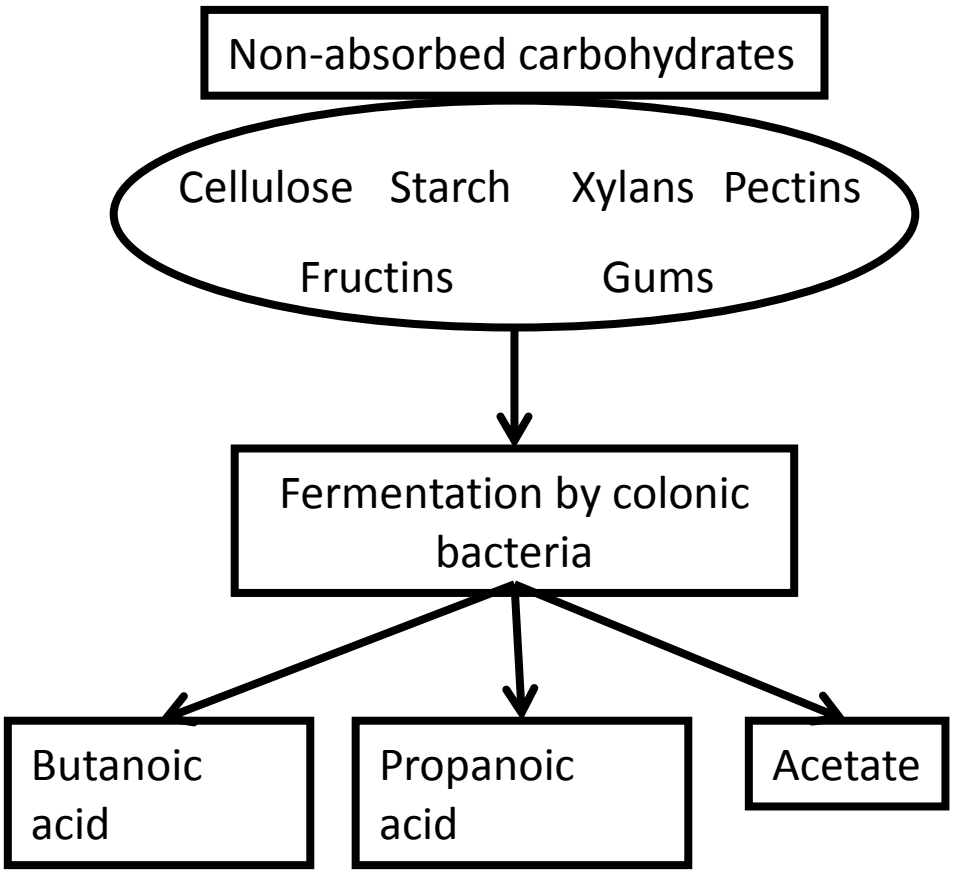
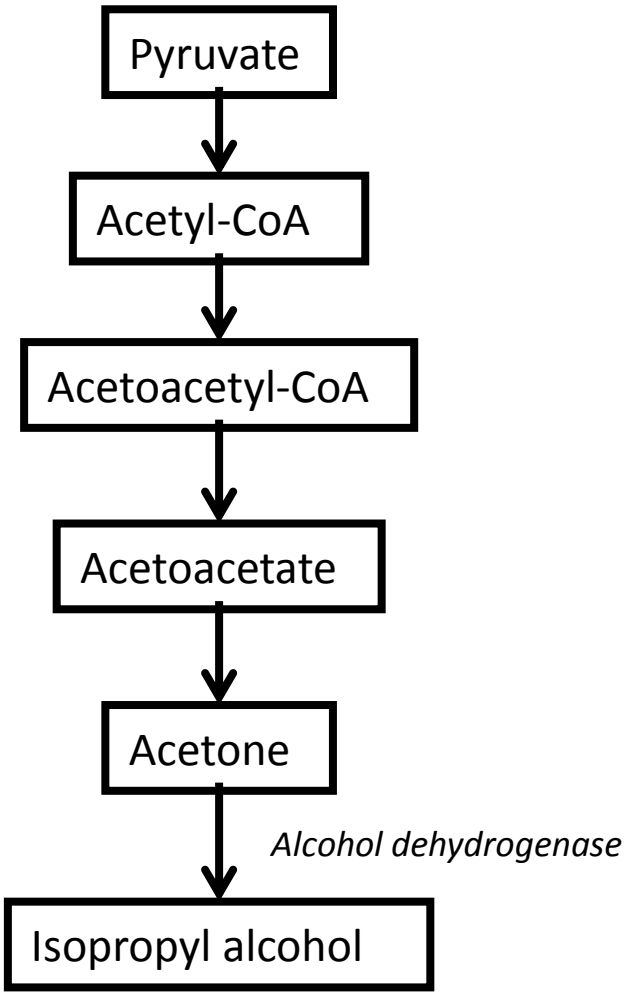


**Figure 52: Pathway for the production of isopropyl alcohol from pyruvate in the human colon. Adapted from associated KEGG pathway.**




Chapter 4 describes the VOCs identified in the headspace gas of the culture medium of *Fusobacterium nucleatum* and *Campylobacter showae*. Both of these bacteria have been implicated in the development of colorectal cancer, via the driver-passenger model[343]. In this study *F. nucleatum* was identified as a net producer of VOCs, and was specifically found to be a net producer of isopropyl alcohol and butanoic acid. *F. nucleatum* has been demonstrated to be over-represented in colorectal cancer[253,299,300,352,353]. Invasive strains of *Fusobacterium nucleatum* accelerate the onset of colonic tumours and drive the transition to a pro-inflammatory microenvironment that is conducive to colorectal tumourigenesis, promoting ongoing tumour growth[268,300]. As

previously described isopropyl alcohol is a recognised carcinogen. Butanoic acid has been shown to promote ongoing tumour growth once it has developed[360] thereby potentially explaining why there is an increase in isopropyl alcohol and butanoic acid based compounds in those with colorectal cancer. *Enterobacteriaceae* have been implicated in the development of colorectal cancer via the production of DNA damaging genotoxins and may thereby actively contribute to the accumulation of mutations that characterize the adenoma-carcinoma sequence[253,373]. *Serratia* spp belong to this family and have been demonstrated to produce isopropyl alcohol[374,375], thus providing another potential source of the VOCs seen within our samples.

It appears from our data that the "signal" from the neoplastic disease becomes stronger as the disease burden increases, either in terms of the size, the number of adenomatous polyps or, as the adenomatous polyp/s become established, carcinoma. This seems most likely to be a product of increasing surface area and overall tissue bulk,but both of these factors will alter the ratio between normal (non-neoplastic) colonic mucosa and abnormal (neoplastic) mucosa, making the VOCs signal more detectable. The associated dysbiosis would potentially become more pronounced as the disease burden increases. A number of studies have examined the composition of the microbiota in patients with colorectal cancer[253,299,300,343,376]. Along with *F. nucleatum* other bacterial species have been noted to differ significantly in both advanced adenomatous disease and colorectal cancer[353]. Wu *et al* described the 16S rRNA patterns seen in 19 patients with colorectal cancer and 20 healthy controls. They observed significant elevation of several bacterial groups, such as *Bacteroides* and *Fusobacterium* species in the colorectal group. Furthermore, there was a positive correlation between *Bacteroides* prevalence and colorectal staging (TNM classification), suggesting increasing dysbiosis with tumour progression[377]. The presence of *F. nucleatum* in colorectal cancer tissue has also been noted in more advanced colorectal cancer, particularly those with lymph node metastasis, supporting the positive correlation[299,377].

### 5.5.1 Conclusion

VOC analysis has a superior diagnostic ability for the identification of colorectal adenocarcinoma, when compared to other faecal based biomarkers, including those currently employed in UK population based screening. It also appears to be superior in the identification of higher risk premalignant adenomatous disease.

# Chapter 6

# An investigation of volatile organic compounds emitted from urine as a biomarker for colorectal neoplasia.

## 6.1 Introduction

Greater levels of population participation in colorectal cancer screening programmes are associated with reduced mortality and greater cost-effectiveness[102]. Worldwide and in the UK there is a large degree of variation in uptake of screening and completion of the stool based testing[103]. There are complex issues behind adherence and uptake of screening programmes, such factors have been reported to include those factors specific to the tests themselves, such as embarrassment and reluctance to handle stool[104]. Urine testing has been reported to be more socially acceptable and therefore has the potential to improve compliance and participation in bowel cancer screening programme.

### 6.1.1 Aim

The aim of this study was to determine if the headspace volatile organic compounds (VOCs) detected in urine samples from patients with neoplastic and no neoplasia in the colon was different, allowing for the assessment of utility of faecal VOCs as a biomarker for the diagnosis and/or monitoring of colorectal neoplasia.

### 6.1.2 Method

SPME headspace extraction followed by GCMS was used to assess VOCs found in the urine of symptomatic and Bowel Cancer Screening Programme (BCSP) patients undergoing colonoscopy. Four millilitres of urine that had been subjected to 24 hours of freeze drying was used and the subsequent diagnosis classified as adenoma, cancer or non-neoplastic, after the colonoscopy. A full description of the methodology can be found in Chapter 2.

## 6.2 Results

### 6.2.1 Cohort demographics

Eighty six recruited patients supplied urine samples. The mean age of the entire cohort was 65 years, with no difference in the mean age of those with and without colonic neoplasia. Fifty seven percent of patients were male.

**Table 33: Demographics and indication for colonoscopy.**

| | **All (%)** | **No-neoplasia (%)** | **Neoplasia (%)** |
|---|---|---|---|
| **n=** | 86 | 38 (44) | 48 (56) |
| **Mean age** | 65 | 65 | 65 |
| Gender | | | |
| **Male** | 51 (60) | 19 (50) | 32 (67) |
| **Female** | 35 (40) | 19 (50) | 16 (33) |
| Indication | | | |
| **IDA** | 23 (27) | 16 (42) | 7 (14) |
| **Surveillance** | 34 (40) | 7 (18) | 27 (56) |
| **BCSP** | 10 (11) | 2 (5) | 8 (17) |
| **CIBH** | 15 (17) | 10 (26) | 5 (10) |
| **Other** | 4 (5) | 3 (7) | 1 (3) |

There were twice as many males with neoplasia than females, the majority of those with neoplasia were referred for surveillance following a previous diagnosis of colonic neoplasia or family history of colorectal cancer. Apart from those subject to screening or surveillance, those referred with iron deficiency anaemia (IDA) were the most numerous, 14% of those found to have neoplasia were referred with IDA.

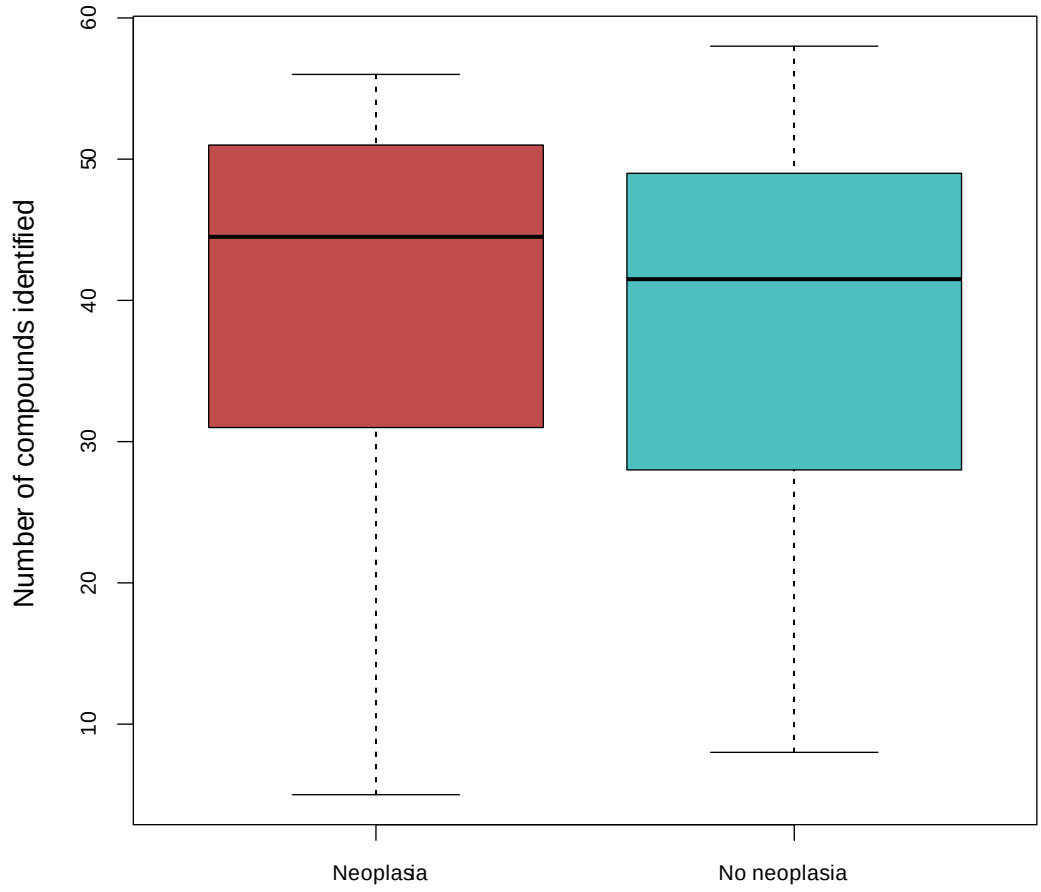
### 6.2.2 VOC Characterisation

A total of 134 compounds were identified across all the samples. There was no significant difference the number of VOCs identified between those with and without neoplasia, *p=0.5,* (Table 34).

**Table 34: Table containing the mean number of VOCs identified in those with and without neoplasia. Includes, standard deviation (SD), standard error of mean (SEM) and coefficient of variation (CV).**

| | **No neoplasia** | **Neoplasia** |
|---|---|---|
| **Mean number of VOCs** | 37.8 | 39.8 |
| **SD** | 14.8 | 14.2 |
| **SEM** | 2.4 | 2.1 |
| **CV** | 0.4 | 0.4 |

**Figure 53: Box and whisker plot demonstrating the number of VOCs identified in the urine of those with and without colonic neoplasia.**




### 6.2.3 Comparisons and biomarker utility assessment

The standard procedures described in Chapter 2, along with the online tool Metaboanalyst, were used to perform the statistical analysis and biomarker assessment. This was able to provide analysis of prevalence and abundance. Of those with neoplasia at colonoscopy, further sub-analysis was also performed according to the size and number of adenomatous polyps identified. Overall the 3 comparisons were made:

- No neoplasia and neoplasia
- No neoplasia and adenoma >1cm in size

- No neoplasia and > 4 individual adenomas of any size

There was no significant difference in the number of VOCs identified between the three groups. When comparison of prevalence was made, only the comparison of no neoplasia and adenoma >1cm demonstrated any significant difference. These compounds were pyrazine, 2,3-dimethyl- (p=0.01) and 6-hepten-3-one, 4-methyl- *(p=0.04)*. No compounds were seen to be significantly different when assessed for abundance across the 3 different comparisons. Principal component analysis and partial least squared discriminant analysis failed to show any meaningful separation. ROC analysis looking at diagnostic utility failed to demonstrate an AUROC of >0.65 across any of the three comparisons.

**6.3 Discussion**

This pilot study failed to demonstrate the diagnostic utility of VOCs emitted from urine as a biomarker for colorectal neoplasia. Across the cohort we were able to detect and measure a significant number of VOCs. There are a number of reasons why this lack of diagnostic utility may have been seen. When comparing the surface area of a single adenomatous polyp to that of the remaining colonic mucosa in which it lies, it is easy to see how small an impact on the overall VOC emission this may have. Other studies that have suggested a diagnostic utility looked specifically at patients with established colorectal cancer, those that examined adenomatous polyps saw a fall in diagnostic ability. Moreover many of these were not assessing urine but faeces[378]. There is clear evidence that VOCs are able to differentiate, IBD, IBS and infective diarrhoea, but these disease states involve larger proportions of the colon than neoplastic disease[379,380]. This is supported by our only finding a significant difference when there was a single adenoma > 1cm.

The production of VOC within the colon is predominately from colonic bacteria. As such, they represent the complex interaction of colonic cells, human gut microflora and invading pathogens[380]. These have been shown to be measurable in urine[309], thought to be possible as a result of gut permeability and the gut-liver axis, a feature which can vary between disease states[381]. This has been shown to increase with colonic neoplasia, with increased permeability by the formation of advanced adenomatous polyps and more so with established colorectal cancer[382]. There was only a relatively small number of patients in this study who had established colorectal cancer or advanced neoplasia. This may explain the lack of diagnostic utility seen. It may

also again be related to surface area and the ratio between normal and abnormal tissue, meaning that there is much greater tissue, with normal permeability, thus, not allowing for differentiation on the basis of identified VOCs. The patients included in this study were recruited via the BCSP and also symptomatic patients attending for colonoscopy. Therefore, of the 48 non-neoplastic patients, a number had other colonic pathology, including Crohn's disease, ulcerative colitis and diverticulosis. The potential for increased gut permeability in the control group was enhanced, suggesting a potential mode for VOC presence in the urine being similar to that seen in neoplasia.

Dietary intake has a significant effect upon VOC production, particularly fermentation and the production of short chain fatty acids[277]. All the patients included in this study provided their samples within 48 hours of their intended colonoscopy. They were all advised to follow a pre-colonoscopy diet, as is the standard practice therefore, there is the potential for similar VOC production as their diets were standardised. The composition of the microbiome can change according to disease states, but if the patients were consuming a similar diet then it has the potential to homogenise the VOC pattern. The dietary modification itself also has the potential to alter the gut microbiome composite[383]. Pre-colonoscopy diets advise the patients to stop consuming items that are high in fibre, such as fruits and cereals, and to increase their fluid intake. This will significantly reduce the production of SCFA such as butyrate, which, in turn, can have an impact upon the microbiome and overall VOC production[15].

When assessing VOCs in the faeces of these patients a series of differences was identified, this was not the case when assessing the VOCs in urine. Many of the VOCs identified, including those deemed significantly different, originate from colonic fermentation e.g. SCFAs. These compounds are promptly absorbed within the colon. They are metabolised in 3 main areas, firstly by the colonic epithelium as a major source of energy. Secondly, liver cells metabolise residual SCFAs for gluconeogenesis. Finally, skeletal muscle tissue can metabolise SCFAs as an energy source. Residual SCFAs are excreted in the faeces[384]therefore renal excretion and urinary detection is very limited. Patients were asked to provide urine samples up to 48 hours before their intended colonoscopy, it is feasible that VOCs were lost from the samples during this time. Urine samples collected as part of the HCC study were frozen within in maximum of 2-3 hours of being produced, potentially preserving VOCs in these samples and allowing for improved detection.

**6.4 Conclusion**

Using the methodology applied to this cohort of samples, there does not appear to be a role for the analysis of VOCs emitted from urine as a biomarker for colorectal neoplasia.

# Chapter 7

# Correlation between faecal tumour M2 pyruvate kinase and colonoscopy for the detection of adenomatous neoplasia in a secondary care cohort

## 7.1 Introduction

There is a well-described relationship between adenoma detection rates and future mortality from colorectal cancer. Therefore, detecting and removing adenomatous polyps is a pertinent medical issue. Colorectal cancer remains a leading cause of mortality and morbidity worldwide. Colorectal cancer screening programmes have been shown to reduce this mortality: many using faecal occult blood testing (FOBt) prior to colonoscopy as the initial screening tool. Other faecal-based diagnostic tools have been proposed, including, tumour M2-pyruvate kinase (tM2-PK).

### 7.1.1 Aim

In the current study, we have used this assay in a novel, secondary care cohort of patients derived from (1) the England Bowel Cancer Screening Programme and (2) symptomatic patients in an attempt to determine whether this assay has a role in diverse settings or as an adjunct to existing FOBt based screening.

### 7.1.2 Method

Patients undergoing colonoscopy in our centre were eligible for inclusion in this prospective study. Patients provided faecal samples immediately prior to bowel preparation. Faecal tM2-PK concentrations were measured using a proprietary ELISA by an investigator blinded to the patients' diagnoses. Sensitivity, specificity, positive predictive value, negative predictive values and ROC analyses were calculated. Kruskal-Wallis and Dunn's post-hoc analyses were applied to numerical data.

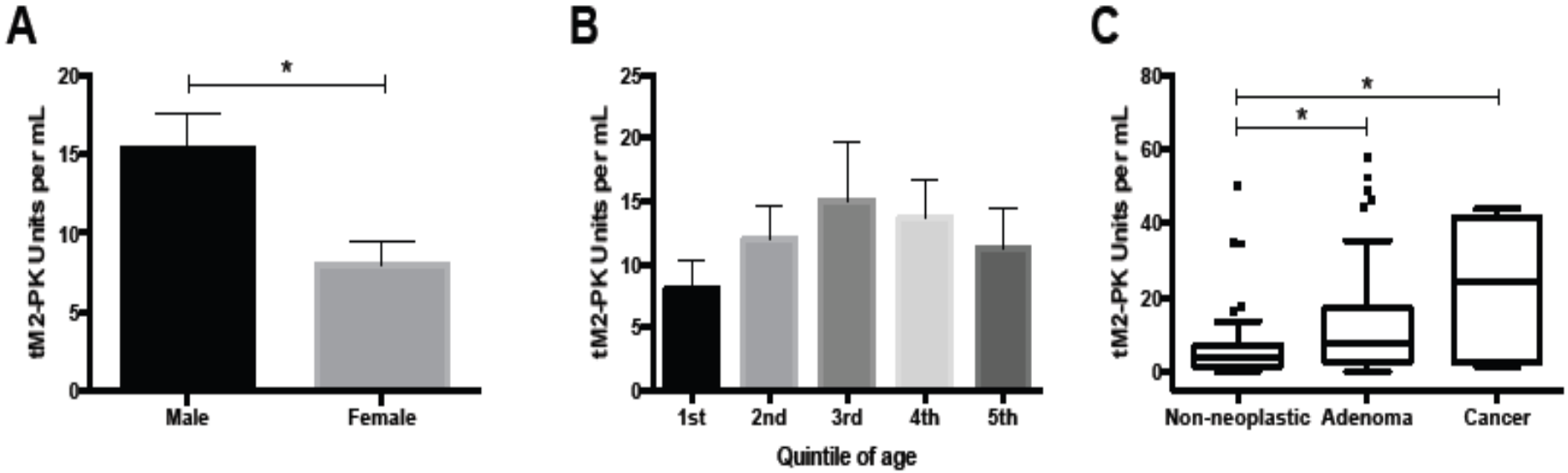
## 7.2 Results

### 7.2.1 Patient demographics

Ninety six patients returned samples within the allotted time. Thirty six patients were participating in the BCSP and, therefore, had a prior, positive g-FOBt. The remaining 60 were symptomatic patients outside the BCSP and had an unknown FOBt status: the indications for their procedures were iron deficiency anaemia (n=18), change in bowel habit (n=15), surveillance or intervention for known polyps (n=20), family history of colorectal cancer and abnormal radiology (n=7). 52 of the cohort were male and 44 female. The median age was 68 years. Faecal tM2-PK levels were significantly higher in males (15.37 U/mL (95% CI 10.8-19.8) vs 8.034 U/mL (95% CI 5.01-10.9), $p$=0.01 by 2-tailed Student's $t$-test) (Figure 54A). To characterise whether patient's age influenced faecal tM2-PK

concentration, the cohort was divided into quintiles according to their age at the time of sampling. No significant differences in mean faecal tM2-PK were identified between any two quintiles (Kruskal-Wallis 1-way ANOVA and Dunn's *post-hoc* analysis; Figure 54B).

**Figure 54: A) Bar chart representing concentration of tM2-PK according to gender of patient (*p<.05, by Student's t-test). B) Bar chart representing concentration of tM2-PK according to age of patients separated by quintiles. C) Box and whisker plot demonstrating tM2-PK concentration stratified by diagnosis at colonoscopy. (*p=≤0.05, by Kruskal-Wallis 1-way ANOVA and Dunns post-hoc analysis).**



### 7.2.2 Correlation of faecal tM2-PK and the progression of colonic neoplasia

Median faecal tM2-PK concentration in individuals with non-neoplastic colonoscopy results was 3.8 U/mL, in those with either adenomatous or malignant disease the concentration was significantly greater (median 7.7 U/mL and 24.4 U/mL, respectively: $p$=0.01 by Kruskal Wallis 1-way ANOVA and Dunn's post-hoc analysis, Figure 54C).

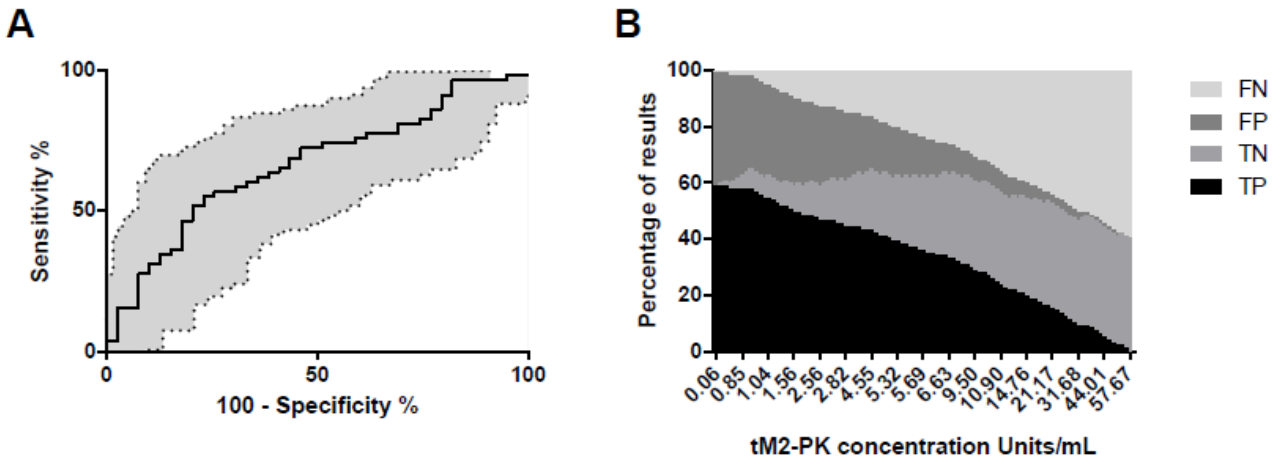
### 7.2.3 Assessment of diagnostic ability of faecal tM2-PK for colorectal neoplasia

To assess the suitability of the faecal tM2-PK ELISA as a screening test in our cohort, we categorised patients into those with non-neoplastic disease and those with a burden of neoplastic disease by colonoscopy. We performed an ROC analysis on the basis of these two groups and showed that the assay may have limited overall utility with an area under the ROC curve of 0.66 ($p$=0.006) (Figure 59A).

The ELISA manufacturer proposed a cut-off value of 4.0 U/mL. Adopting this as a cut-off to define a positive tM2-PK test, the assay's sensitivity to detect adenoma or carcinoma was 72.4 % (95 % CI; 59.1-83.3 %), specificity 48.7 % (95 % CI; 32.4-65.2 %), positive predictive value 67.7 % (95 % CI; 54.7-79.2 %) and negative predictive value 36.7 % (95 % CI; 36.7-71.2 %). Application of McNemar's test did not show a significant difference between the tM2-PK results and the presence of neoplasia at colonoscopy ($p$=0.73), suggesting a limited relationship between the faecal tM2-PK value and findings at colonoscopy. There was a false positive rate of 20.6 % across the 96 samples. In total, 37 % of samples were incorrectly categorised by tM2-PK testing.

The same statistical analysis was performed using a series of different threshold levels (Table 35 and Figure 55B). The optimal cut-off value identified by the highest McNemar's test was at 4.8: this gave a sensitivity of 69 % and a specificity of 56 %.

**Figure 55: A) ROC curve showing utility of faecal tM2-PK ELISA for detection of adenomatous disease in this cohort, shaded area represents 95 % confidence intervals. B) Surface plot demonstrating the variation in diagnostic accuracy as the diagnostic threshold concentration changes.**



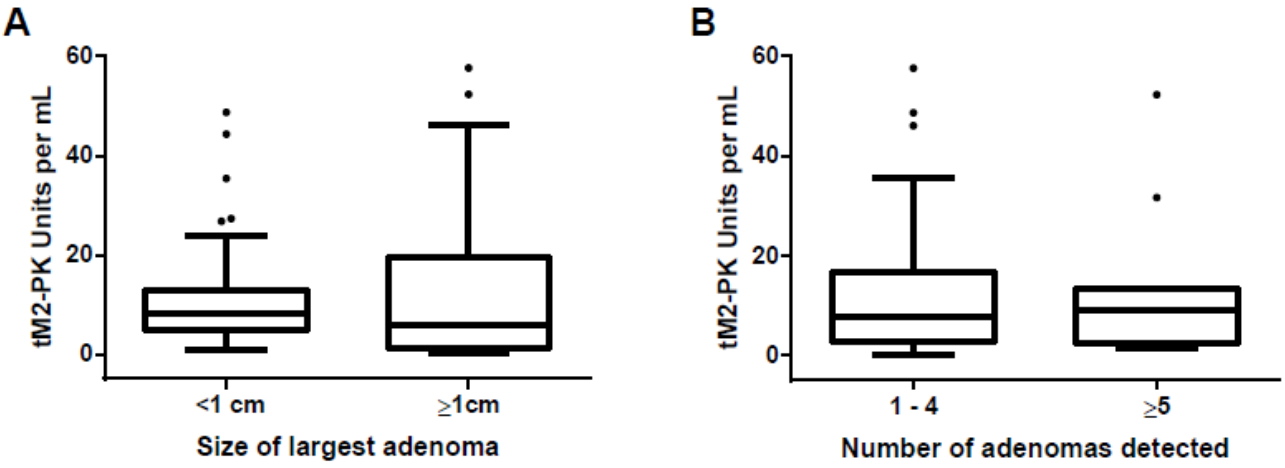

FN: false negative, FP: false positive, TN: true negative, TP: true positive.

**Table 35: Table showing the sensitivity and specificity of different faecal tM2-PK thresholds for the study cohort, with binomial method McNemar values for each of these thresholds.**

| Threshold | Sensitivity % | 95 % CI | Specificity % | 95 % CI | Likelihood ratio | McNemar Binomial p-value |
|---|---|---|---|---|---|---|
| 3 | 75.86 | 62.83 to 86.13 | 41.03 | 25.57 to 57.90 | 1.286 | 0.2559 |
| 3.2 | 74.14 | 60.96 to 84.75 | 41.03 | 25.57 to 57.90 | 1.257 | 0.2559 |
| 3.4 | 74.14 | 60.96 to 84.75 | 43.59 | 27.81 to 60.38 | 1.314 | 0.324 |
| 3.6 | 74.14 | 60.96 to 84.75 | 43.59 | 27.81 to 60.38 | 1.314 | 0.324 |
| 3.8 | 74.14 | 60.96 to 84.75 | 48.72 | 32.42 to 65.22 | 1.446 | 0.6177 |
| 4 | 72.41 | 59.10 to 83.34 | 48.72 | 32.42 to 65.22 | 1.412 | 0.7359 |
| 4.2 | 72.41 | 59.10 to 83.34 | 48.72 | 32.42 to 65.22 | 1.412 | 0.7359 |
| 4.4 | 72.41 | 59.10 to 83.34 | 51.28 | 34.78 to 67.58 | 1.486 | 0.7359 |
| 4.6 | 72.41 | 59.10 to 83.34 | 53.85 | 37.18 to 69.91 | 1.569 | 0.8642 |
| 4.8 | 68.97 | 55.46 to 80.46 | 56.41 | 39.62 to 72.19 | 1.582 | 0.8679 |
| 5.0 | 67.24 | 53.66 to 78.99 | 56.41 | 39.62 to 72.19 | 1.543 | 0.7428 |
| 5.2 | 65.52 | 51.88 to 77.51 | 56.41 | 39.62 to 72.19 | 1.503 | 0.7428 |

Of the 50 patients with adenomatous disease 22 were found to have a single polyp greater than 1cm in size. The median faecal tM2-PK reading for this group was 5.7 U/mL (interquartile range 1.3-19.5). Of these patients, 59 % (n=13) had a faecal tM2-PK result above the 4.0 U/mL cut-off, representing a true positive result by the assay manufacturer's criteria. There was no significant difference in the tM2-PK concentration detectable in the faeces of patients with polyps >1cm or <1cm, *p*=0.63 (Figure 60A). There was also no significant difference in the tM2-PK concentration in faeces of those with <5 or ≥5 adenomatous polyps (Figure 56B).

**Figure 56: A) Box and whisker plot showing faecal tM2-PK concentration stratified by size of largest adenoma. B) Box and whisker plot of faecal tM2-PK concentration stratified by the number of adenomas identified in each patient.**

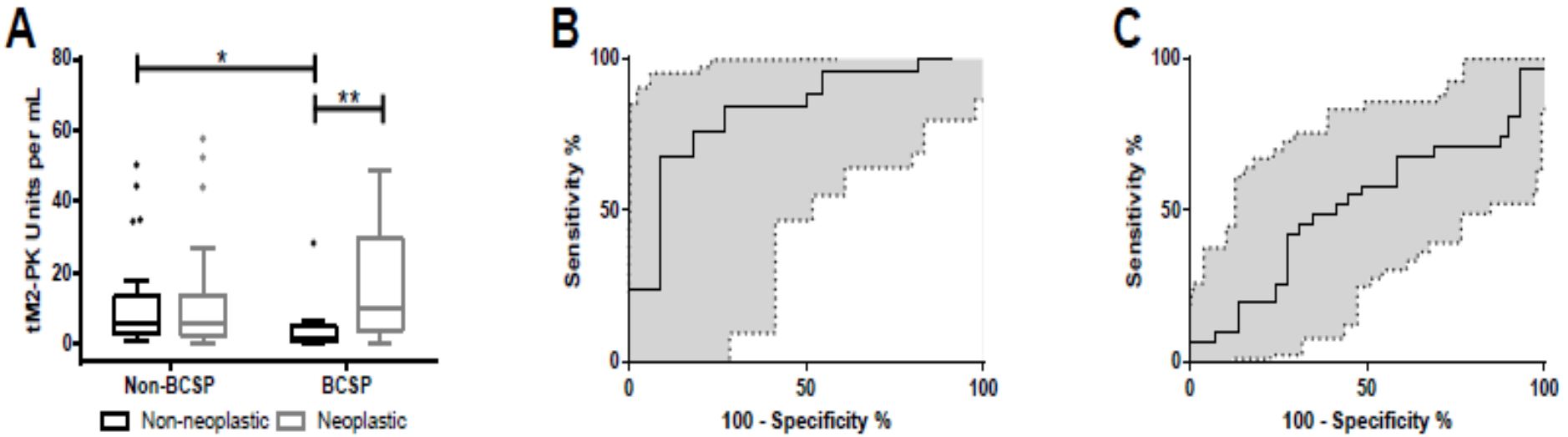



### 7.2.4 Assessment of faecal tM2-PK as adjunct to gFOBt in BCSP

When we isolated those individuals (n=36) referred for colonoscopy, as part of the BCSP, those with neoplasia (n=25) were found to have significantly greater faecal tM2-PKconcentration than those without neoplastic pathology (Figure 57A, *p*=0.03). A macroscopically normal colonoscopy was seen in all BCSP patients who did not otherwise have neoplastic disease (n=11). ROC analysis of the BCSP patients alone demonstrated an AUROC of 0.82 (*p*=0.002) (Figure 61B). Sixty one percent of patients with a prior positive fFBOt also had a positive tM2-PK. Of the BCSP with no neoplasia, 72% had a negative tM2-PK, whilst 76% of those with neoplasia had a positive tM2-PK.

This difference was not observed in the non-BCSP group; 29 of whom were classified as non-neoplastic, of these 7 had IBD and/or diverticulosis. Within this subset, 6 had elevated (false positive) tM2-PK concentration. The remaining 22 subjects had a macroscopically normal colonoscopy, and 11 received a true negative faecal tM2-PK result, whilst 11 had falsely positive results. Within this subset AUROC was not significantly different to 0.5 (AUROC 0.51, $p$=0.87)(Figure 57C).

**Figure 57: A) Box and whisker plot demonstrating faecal tM2-PK concentration in patients recruited via BCSP or non-BCSP route, and with or without adenomatous pathology (*p<0.05, **p<0.01, by Kruskal Wallis 1-way ANOVA and Dunn's post-hoc analysis). B) ROC curve demonstrating the diagnostic accuracy of faecal tM2-PK ELISA in patients recruited following colonoscopy referral from the BCSP. C) ROC curve demonstrating the diagnostic accuracy of faecal tM2-PK ELISA in patients not recruited via the BCSP. Shaded areas represent 95 % CI.**



In our secondary care cohort, the faecal tM2-PK assay appears to have a sensitivity that is comparable to, and in some instances, superior to existing faecal markers used as part of colorectal cancer screening programmes. The present study demonstrates a superior sensitivity to that associated with g-FOBt. However, the main issues that will prevent its independent use are its comparatively low specificity, high false positive rate and time constraints for sample collection. Its use with existing FOBt methods does, however, seem appealing. Combining the two methods appears to rectify some of the flaws of each, and could improve the accuracy of an initial faecal biomarker screening tool. One role may be the sequential use of gFOBt followed by tM2-PK. However, the introduction of this type of screening would represent additional time and financial burdens that may constrain the use of these investigations. Ultimately, in order to justify their use, a reduction in mortality from colorectal cancer with such screening tools would need to be demonstrated.

**7.3 Discussion**

The faecal tM2-PK concentration observed in this study demonstrated an increasing trend as pathology progressed along the adenoma-carcinoma pathway, supporting potential utility for the assay to distinguish between absence and presence of neoplastic disease.

Our cohort contained more males, with a median age of 68, and was therefore, representative of a potential colorectal cancer population. The faecal tM2-PK concentration was higher in males, with age apparently having no impact on the abundance of this enzyme. In a clinical setting, this lack of variation with age could be deemed beneficial.

Based on the manufacturer's threshold value of 4 U/mL, we found that the assay's sensitivity to detect any adenomatous pathology was 72 % with specificity of 48 %. This is consistent with the findings of a meta-analysis, from 2012, which examined seventeen studies and found the quoted sensitivities, for colorectal cancer, to range from 68-97 % and adenoma detection to range from 28-76 %.Within the meta-analysis several of the studies focused solely on the detection of carcinoma, rather than adenomatous polyps.

To optimise the diagnostic accuracy of a faecal tM2-PK assay ,within a future screening programme, the diagnostic threshold for the assay could be adjusted away from the manufacturer's recommended threshold of 4 U/mL. To understand whether this process may have utility we applied

a series of thresholds to our data set and calculated binomial McNemar statistics to define the threshold with optimal diagnostic utility for our cohort. Using this methodology we demonstrated optimal utility with a threshold of 4.8 U/mL. This led to an improved sensitivity of 69.0 % and specificity of 56.4 % in our cohort. At this threshold 40.2 % of tests yielded true positive results, 22.7% a true negative, whilst 17.5 % were false positive results and 19.5 % were false negative results. One of the obvious frailties of this methodology is that with this statistical approach the consequences of false negative and false positive results are conferred equal weighting. This is clearly not the case in this clinical situation, where the consequences of a false negative result manifestly outweigh those of a false positive.

Using the manufacturer's recommended threshold of 4 U/mL the faecal tM2-PK assay had a lower false negative rate of 16.5 %, but an increased false positive rate of 20.6 % in our cohort. Whilst this threshold could reduce the number of patients with adenomatous disease that were not detected by the faecal tM2-PK assay, it would lead to an increased number of endoscopic procedures that yielded no adenomatous pathology being performed. This false positive rate is substantially higher than that reported by Tonus *et al*[194], but is consistent with a study from 2006 which reported a low specificity with high false positive rate [385], suggesting that it may be within the range expected in a clinical cohort. One study from the USA estimated that a 5–10 % false positive rate during colorectal screening would amount to an expense of US$6 billion if 100 % participation rate is assumed [386]. The current UK Bowel Cancer Screening Program reports a sensitivity of 36.5 % and specificity of 92.2 % for g-FOBt. The literature suggests that if a different screening tool were to be implemented, for example i-FOBt, a higher detection rate could be achieved e.g. a sensitivity of 70.9 % and specificity of 96.3 % [387]. The consequence of a false negative test in bowel cancer screening is grave and, therefore, the lower specificity associated with tM2-PK testing will mean it is unlikely to replace either of these methods in colorectal screening or as a preliminary test in a secondary care based population. False positive results have been reported in inflammatory bowel disease, infective disease and diverticulosis[388,389], this is replicated in our cohort as those with non neoplastic pathology and no prior FOBt had a false positive rate of 50 %.

The combination of FOBt with faecal tM2-PK ELISA has been proposed as a strategy to optimise the sensitivity and specificity for a screening programme. When we segregated our cohort into those recruited from the BCSP and those with other indications for colonoscopy we demonstrated greater diagnostic utility for tM2-PK ELISA in patients recruited from the bowel cancer screening population. This was associated with a significantly lower median tM2-PK concentration in the faeces of patients without adenomatous disease in the BCSP cohort compared to those without adenomatous disease

in the non-BCSP cohort. This suggests that the utility of faecal tM2-PK as a biomarker for colonic adenomatous disease may be limited to the relatively homogenous asymptomatic screening cohort, rather than the more heterogenous group of symptomatic patients.

A study from 2014 assessed the utility of g-FOBt with faecal tM2-PK ELISA prior to potential colonoscopy. 1800 individuals were invited to participate, with a response rate of 54 % (n=978), culminating in over 800 analysable samples being returned and 186 colonoscopies being performed. Overall, positivity was significantly increased by the addition of tM2-PK (27 %). Had only i-FOBt been tested, 77 % (n=189) fewer patients would have had a positive faecal test and been identified for a screening colonoscopy and 70 % (n=35) fewer patients would have had polyps detected and removed[390]. This study also reported that 10 % (n=99) of tM2-PK samples could not be analysed, because of a lapse in the 48 hour window required for its processing. This is clearly an issue for a postal based screening programme as is currently in operation in the UK. A further study from Italy assessed the impact of combined testing on 280 samples. For colorectal cancer detection, i-FOBt was the test with the highest specificity and positive predictive value (0.89 and 0.53), whereas tM2-PK had the highest sensitivity and negative predictive value (0.87 and 0.96). It demonstrated a good ability of the combined test to identify colorectal cancer, with patients showing positivity to both markers, the risk of cancer was as high as 79 % [391]. Other assays have been assessed in combination with FOBt, including faecal calprotectin, which demonstrated sensitivity for the detection of colorectal cancer of 79 %, whereas the sensitivity for the combination of i-FOBt and tM2-PK was 93 % [392].

An analysis of 697 patients in the Cleveland Clinic Foundation Adenoma Registry showed that, compared with one to two small adenomas, risk is increased 10-fold after removal of multiple adenomas at least one of which is larger than 1 cm [393], and adenomas of >1cm diameter identify individuals at increased risk of colorectal cancer even if removed [394]. It is therefore clinically important to identify polyps of >1cm diameter. With increasing size, one might predict increased abundance of tM2-PK, however within our cohort this was not the case.

There is a similar increase in risk for those with ≥5 individual polyps, even if they are all <1cm in size. The apparent utility of the tM2-PK assay to detect polyps irrelevant of size makes it appealing for use in a screening or surveillance programme. Another beneficial element of this finding is that polyps <1cm do not tend to bleed and, therefore, may be missed by FOBt [395]. The tM2-PK assay does not have this problem as it does not rely on the presence of blood and, therefore, has the potential to detect non-bleeding adenomatous disease. Moreover, in our cohort no difference in the abundance of tM2-PK was identified in those with 5 or more adenomas, compared to those with fewer.

The application of tM2-PK shows potentially beneficial outcomes when used in a homogenous clinical group, such as the BCSP. The interference from a heterogeneous population, such as our secondary care cohort appears to impact upon the diagnostic utility of faecal tM2-PK analysis. A dedicated randomised controlled study would be required to explore this potential combination further. We must also concede that the relatively small number of patients included in this study is a limitation.

### 7.3.1 Conclusion

In our secondary care cohort, the faecal tM2-PK assay appears to have a sensitivity that is comparable to, and in some instances, superior to existing faecal markers used as part of colorectal cancer screening programmes. The present study demonstrates a superior sensitivity to that associated with g-FOBt. However, the main issues that will prevent its independent use are its comparatively low specificity, high false positive rate and time constraints for sample collection. Its use with existing FOBt methods does however seem appealing. Combining the two methods appears to rectify some of the flaws of each, and could improve the accuracy of an initial faecal biomarker screening tool. One role may be the sequential use of gFOBt followed by tM2-PK. However, the introduction of this type of screening would represent additional time and financial burdens that may constrain the use of these investigations. Ultimately, in order to justify their use, a reduction in mortality from colorectal cancer with such screening tools would need to be demonstrated.

# Chapter 8

# Concluding discussion and future work

## 8.1 Concluding discussion

The UK incidence of both colorectal cancer and HCC is increasing in the UK. Screening and early diagnosis of gastrointestinal malignancies has the ability to improve outcomes by reducing morbidity and mortality[5,129]. In the UK screening for HCC and colorectal cancer is currently adopted. There is clear evidence that the UK BCSP has reduced colorectal cancer associated mortality, namely via the identification of pre-malignant adenomatous polyps, leading to their removal and thus preventing formation of adenocarcinoma. The evidence for such an impact of HCC screening in cirrhotic patients is less robust, but is supported by national and international guidance[127,129]. Both programmes have areas that could be improved. For the BCSP, the FOBt used as a pre-screen to colonoscopy leads to a high false positive rate and thus potentially unnecessary colonoscopy. This has significant financial implications for the screening programme and NHS. Screening for HCC is currently based upon periodic USS assessment. Therefore, the process is not employing a point of care test or pre-screening test prior to the diagnostic test. As evidenced by my work and other work in the published literature, VOCs appear to have the potential to supplement both programmes in order to improve the overall diagnostic accuracy and thus overall impact, whilst limiting expenditure[306,311,312,365].

A point of care test either used in general practice or as part of population based screening, conducted via a sensor based technology appears to be the most appropriate application of VOCs as a biomarker for gastrointestinal neoplasia[396]. For both colorectal cancer and HCC such technology would be best placed to select those required to undergo further investigations, i.e. colonoscopy for colorectal cancer and USS for HCC. Using such “pre-screen” methods with improved sensitivity would limit the number of false positives that lead to unnecessary costly investigation. Obviously this should not be at the expense of inferior specificity as the consequence of a false negative is a missed cancer and therefore could have grave implications.

Other faecal based assays are available for the diagnosis of colorectal cancer, for example the faecal tM2-PK assay that is described in chapter 7. The main issues that will prevent its independent use are its comparatively low specificity, high false positive rate and time constraints for sample collection. Its use with existing FOBt methods does however seem appealing. Combining the two methods appears to rectify some of the flaws of each, and could improve the accuracy of an initial faecal biomarker screening tool. One role may be the sequential use of gFOBt followed by tM2-PK. VOCs appear to have a superior performance when compared to tM2-PK for the diagnosis of colorectal cancer.

Ultimately, in order to justify the use of either VOCs or other screening tool a reduction in mortality and morbidity from colorectal cancer or HCC will have to be demonstrated.

The genetic abnormalities associated with the development of colorectal cancer are well described. Less is known about the association between the host microbiota and the development of colorectal cancer. It can be difficult to prove a causative role for specific bacteria, but attempts have been made through the description of such models as the bacterial driver-passenger model of colorectal cancer[343]. This model suggests that as the colonic mucosa becomes more neoplastic there is preferential selection of passenger bacteria that overpopulate the abnormal tissue. Their actions, including their metabolic productions, promote the ongoing existence and growth of the neoplastic tissue[350,353].

There are a number of limitations within the studies described that should be highlighted. Regarding the investigation of VOCs emitted from stool as a biomarker for colorectal neoplasia the number of cancer samples were limited along with the clinical information relating to them. Ideally there would have been a larger number of cancer samples across all 3 regions of recruitment. This would allow for a better assessment of inter-regional variability of VOCs. Staging of the cancer, either radiologically or histologically would have improved analysis. This would have allowed the separation of potentially curable disease, much like that explored in the VOCs emitted from urine as a biomarker for HCC study. A further point of improvement would have been to record the anatomical location of each cancer or polyp in order to ascertain if there is a difference between left and right sided disease. This is of particular importance as the Bowel Scope screening programme is being rolled out across England.

Additional testing of the gFOBt status of all the non-BCSP samples would allow for further analysis to be performed. This would allow for a direct comparison of the diagnostic accuracy of VOCs and gFOBt in order to determine if VOCs are superior. It would also further explore the whether the two approaches could be synergistic in the identification of neoplasia. This would also allow for further exploration of the results seen in chapter 7.

Recruitment for the HCC study took place in Liverpool only. This was across two centres but still within a single geographic region, this could be considered as a weakness. Moreover the majority of HCC samples were recruited from a single centre, as this acts as the regional referral hospital. As discussed in chapter 3 the VOCs emitted from urine can be influenced by a number of factors including ethnicity and diet. Therefore to improve this element future work should aim to recruit from a number of regions within the UK or even on an international scale. It should also aim to

match its recruitment to the ethic profile of the UK, according to census data. This should also be in conjunction with consideration to recruit pathology equally across centres in order to account of differences in disease/patient management.

**8.2 Future work**

There are a number of avenues to explore as a result of the work discussed. This includes larger scale duplication studies, mechanistic exploration of the associated microbiota and origins of the VOCs identified and application of the methodology to other pertinent medical issues.

A larger scale study would certainly be required for both the faecal colorectal cancer VOC study and the urinary HCC VOC study. The larger colorectal study, would involve a number of sites across the UK in order to account for any geographical variability in diet and ethnic composition of the populations. It would also involved recruitment of a much larger cohort, allowing for sub-analysis of adenoma size, histological appearance and anatomical location of colorectal cancer and adenomas. FOBt could also be conducted on all stool in order to directly compare the performance of VOCs to FOBt as a pre-screen to colonoscopy and whether combining the two could improve selection for colonoscopy. The discovery study described in chapter 5 identified a 3 VOC biomarker panel. A larger scale study would act to explore/confirm the application of this panel. This could be through a tailored GCMS methodology, or via a new sensor technology, such as Odoreader© or new in-house collaborative sensor technology specifically designed to only identify the 3 VOCs. Clearly the origins of significant VOCs needs exploration and explanation. Bacterial genetic sequencing technologies can be employed in order to detect difference in the microbiota associated with the presence of adenoma and/or carcinoma, this can be performed on the stool provided. Metabolic pathways and VOC production by specific pathogens can then be explored, as per chapter 4. VOC analysis of colorectal cancer organoid culture may also allow for the assessment of the source of significant VOCs.

A larger, multi-centred study for the use of urinary VOCs for the diagnosis of HCC would address geographical issues identified in the study conducted. It would also allow for an increase in the range and number of underlying aetiologies, thus being more representative of a screening population. Once again a panel of VOCs can be better explored with other sensor technologies. It would also allow for the assessment of the model described, particularly those identifying as diagnostic for BCLC curative HCC. Coupled to this larger study, VOC analysis of HCC tissue from resected tumours could be performed, allowing for the explanation of the origin of such compounds.

A prospective study, whereby urinary VOCs are coupled to USS surveillance would allow the assessment of the use of VOCs as a pre-screen tool to USS based screening.

The methodologies used to explore colorectal cancer and HCC could be applied to other gastrointestinal issues. One such issue is the early identification of intestinal anastomotic leak following surgical resection. Patients undergoing intestinal resection with a primary anastomosis are at risk of anastomotic breakdown and leak of luminal content into the peritoneum. This has been reported to be in the region of 5% of cases[397]. Particularly in oncological procedures, this results in worse outcomes and a reduced 5 year survival[398]. Typically in the immediate post operative period surgical drains are placed and remain in situ for 2-3 days post op. As luminal content leaks into the peritoneum, so too does the gut microbiota, one could assume that the associated VOC profile in the drain fluid would alter to represent the presence of such bacteria. Periodic VOC analysis of the drain fluid could be used to identify those who have leaked or ideally, in whom a leak is imminent, thus requiring further surgical intervention before intra-abdominal sepsis occurs.

**References**


1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013 Apr;49(6):1374–403.

2. CRUK. http://publications.cancerresearchuk.org/downloads/Product/CS_KF_BOWEL.pdf. 2014;

3. Rees CJ, Bevan R. The National Health Service Bowel Cancer Screening Program: the early years. Expert Rev Gastroenterol Hepatol. 2013 Jul;7(5):421–37.

4. Trueman P, Lowson K, Chaplin S, Wright D, Duffy S, Officer I. Bowel Cancer Services : Costs and Benefits. 2007.

5. Logan RF a, Patnick J, Nickerson C, Coleman L, Rutter MD, von Wagner C. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut. 2012 Oct;61(10):1439–46.

6. Durko L, Malecka-Panas E. Lifestyle Modifications and Colorectal Cancer. Curr Colorectal Cancer Rep. 2014 Jan;10:45–54.

7. OS L. Acquired risk factors for colorectal cancer. Methods Mol Biol. 2009;472:361–72.

8. Burkitt DP. Related disease--related cause? Lancet. 1969;2:1229–1231.

9. Research WCRF/ AI for C. Food, nutrition, physical activity and the prevention of colorectal cancer. 2011.

10. Scharlau D, Borowicki A, Habermann N, Hofmann T, Klenow S, Miene C, Munjal U, Stein K, Glei M. Mechanisms of primary cancer prevention by butyrate and other products formed during gut flora-mediated fermentation of dietary fibre. Mutat Res. 2009;682(1):39–53.

11. Coxhead JM, Williams E a, Mathers JC. DNA mismatch repair status may influence anti-neoplastic effects of butyrate. Biochem Soc Trans. 2005 Aug;33(Pt 4):728–9.

12. Scheppach W, Weiler F. The butyrate story: old wine in new bottles? Curr Opin Clin Nutr Metab Care. 2004 Sep;7(5):563–7.

13. Nakano K, Mizuno T, Sowa Y, Orita T, Yoshino T, Okuyama Y, Fujita T, Ohtani-Fujita N, Matsukawa Y, Tokino T, Yamagishi H, Oka T, Nomura H, Sakai T. Butyrate Activates the WAF1/Cip1 Gene Promoter through Sp1 Sites in a p53-negative Human Colon Cancer Cell Line. J Biol Chem. 1997 Aug 29;272(35):22199–206.

14. Reddy BS, Hirose Y, Cohen LA. Preventive Potential of Wheat Bran Fractions against Experimental Colon Carcinogenesis : Implications for Human Colon Cancer Prevention Preventive Potential of Wheat Bran Fractions against Experimental Colon Carcinogenesis : Implications for Human Colon C. Cancer Res. 2000;4792–7.

15. Canani RB, Costanzo M Di, Leone L, Pedata M, Meli R, Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. World J Gastroenterol. 2011 Mar 28;17(12):1519–28.

16. Bingham SA, Day NE, Luben R et al. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an

observational study. Lancet. 2003;361:1496–1501.

17. Aune D, Chan DSM, Lau R, Vieira R, Greenwood DC, Kampman E, Norat T. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. BMJ. 2011 Jan;343(November):d6617.

18. Asano TK MR. Dietary fibre for the prevention of colorectal adenomas and carcinomas ( Review )-Cochrane Database Syst Rev. 2008 CD003430.

19. Bingham S. The fibre–folate debate in colo-rectal cancer. Proc Nutr Soc. 2007 Mar 7;65(1):19–23.

20. Powers HJ. Interaction among folate, riboflavin, genotype, and cancer, with reference to colorectal and cervical cancer. J Nutr. 2005;135:2960S–2966S.

21. Wang TP, Hsu SH, Feng HC HR. Folate deprivation enhances invasiveness of human colon cancer cells mediated by activation of sonic hedgehog signaling through promoter hypomethylation and cross action with transcription nuclear factor-kappa B pathway. Carcinogenesis. 2012;33(6):1158–1168.

22. Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, McKeown-Eyssen G, Summers RW, Rothstein RI, Burke CA, Snover DC, Church TR, Allen JI, Robertson DJ, Beck GJ, Bond JH, Byers T, Mandel JS, Mott LA, Pearson LH, Barry EL, Rees JR, Marcon N, Sa PPSG. Folic acid for the prevention of colorectal adenomas: a randomized clinical trial. JAMA. 2007;297(21):2351–9.

23. Larsson SC, Orsini N WA. Vitamin B6 and risk of colorectal cancer: a meta-analysis of prospective studies. JAMA. 2010;303(11):1077–83.

24. Chan DS, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E NT. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. PLoS One. 2011;6:e20456.

25. Daniel CR, Cross AJ, Koebnick C SR. Trends in meat consumption in the USA. Public Heal Nutr. 2011;14(4):575–83.

26. Knöbel Y, Weise A, Glei M, Sendt W, Claussen U P-ZB. Ferric iron is genotoxic in non-transformed and preneoplastic human colon cells. Food Chem Toxicol. 2007;45(5):804–11.

27. Kim E, Coelho D, Blachier F. Review of the association between meat consumption and risk of colorectal cancer. Nutr Res. 2013 Dec;33(12):983–94.

28. Toden S, Bird AR, Topping DL CM. Resistant starch attenuates colonic DNA damage induced by higher dietary protein in rats. Nutr Cancer. 2005;51:45–51.

29. Toden S, Bird AR, Topping DL CM. High red meat diets induce greater numbers of colonic DNA double-strand breaks than white meat in rats: attenuation by highamylose maize starch. Carcinogenesis. 2007;28:2355–62.

30. Ollberding NJ, Wilkens LR, Henderson BE, Kolonel LN L, L. M. Meat consumption, heterocyclic amines and colorectal cancer risk: the Multiethnic Cohort Study. Int J Cancer. 2012;131(7):E1125–33.

31. Alexander DD CC. Red meat and colorectal cancer: a critical summary of prospective epidemiologic studies. Obes Rev. 2011;12:e472–93.

32. Leufkens AM, Van Duijnhoven FJ, Siersema PD, Boshuizen HC, Vrieling A, Agudo A, Gram IT, Weiderpass E, Dahm C, Overvad K, Tjønneland A, Olsen A, Boutron-Ruault MC, Clavel-Chapelon F, Morois S, Palli D, Grioni S, Tumino R, Sacerdote C, Mattiello A, Herman B-MH. Cigarette smoking and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition study. Clin Gastroenterol Hepatol. 2011;9(2):137–44.

33. Wong HP, Yu L, Lam EK, Tai EK, Wu WK CC. Nicotine promotes colon tumor growth and angiogenesis through beta-adrenergic activation. Toxicol Sci. 2007;97(2):279–87.

34. Haas SL, Ye W LJ. Alcohol consumption and digestive tract cancer. Curr Opin Clin Nutr Metab Care. 2012;15(5):457–67.

35. Cai S, Li Y, Ding Y, Chen K JM. Alcohol drinking and the risk of colorectal cancer death: a meta-analysis. Eur J cancer Prev. 2014;23:532–9.

36. Cho E, Smith-Warner SA, Ritz J, van den Brandt PA, Colditz GA, Folsom AR, Freudenheim JL, Giovannucci E, Goldbohm RA, Graham S, Holmberg L, Kim DH, Malila N, Miller AB, Pietinen P, Rohan TE, Sellers TA, Speizer FE, Willett WC, Wolk A HD. Alcohol intake and colorectal cancer: a pooled analysis of 8 cohort studies. Ann Intern Med. 2004;140:603–13.

37. Blasiak J, Trzeciak A, Malecka-Panas E, Drzewoski J WM. In vitro genotoxicity of ethanol and acetaldehyde in human lymphocytes and the gastrointestinal tract mucosa cells. Toxicol Vitr. 2000;14(4):287–95.

38. Ben Q, Wang L, Liu J, Qian A, Wang Q YY. Alcohol drinking and the risk of colorectal adenoma: a dose-response meta-analysis. Eur J cancer Prev. 2014;[Epub ahead of print.

39. Ma Y, Yang Y, Wang F, Zhang P, Shi C, Zou Y QH. Obesity and risk of colorectal cancer: a systematic review of prospective studies. PLoS One. 2013;8:e53916.

40. Ning Y, Wang L G EL. A quantitative analysis of body mass index and colorectal cancer: findings from 56 observational studies. Obes Rev. 2010;11(1):19–30.

41. Aleksandrova K, Nimptsch K PT. Obesity and colorectal cancer. Front Biosci (Elite Ed. 2013;1:61–77.

42. Booth FW, Chakravarthy MV, Gordon SE SE. Waging war on physical inactivity: using modern molecular ammunition against an ancient enemy. J Appl Physiol. 2002;93(1):3–30.

43. Kumor A, Daniel P, Pietruczuk M M-PE. Serum leptin, adiponectin, and resistin concentration in colorectal adenoma and carcinoma (CC) patients. Int J Colorectal Dis. 2009;24(3):275–81.

44. Wolin KY, Yan Y, Colditz GA LI. Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer. 2009;100(4):611–6.

45. Wolin KY, Yan Y CG. Physical activity and risk of colon adenoma: a meta-analysis. Br J Cancer. 2011;104(5):882–5.

46. Halle M SM. Physical activity in the prevention and treatment of colorectal carcinoma. Dtsch Aerzteblatt Int. 2009;106(44):722–7.

47. Wiseman M. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Proc Nutr Soc. 2008;67(3):253–6.

48. Robsahm TE, Aagnes B, Hjartåker A, Langseth H, Bray FI LI. Body mass index, physical activity,

and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. Eur J cancer Prev. 2013;22(6):492–505.

49. Drevon CA. Fatty acids and expression of adipokines. Biochim Biophys Acta. 2005;1740(2):287–92.

50. Poulsen HE, Weimann A LS. Methods to detect DNA damage by free radicals: relation to exercise. Proc Nutr Soc. 1999;58(4):1007–14.

51. Innominato PF, Roche VP, Palesh OG, Ulusakarya A, Spiegel D LF. The circadian timing system in clinical oncology. Annu Med. 2014;46(4):191–207.

52. Hrushesky WJ, Lannin D HE. Evidence for an ontogenetic basis for circadian coordination of cancer cell proliferation. J Natl Cancer Inst. 1998;90(19):1480–4.

53. Ko CH TJ. Molecular components of the mammalian circadian clock. Hum Mol Genet. 2006;2:R271-7.

54. Wood PA, Yang X, Taber A, Oh EY, Ansell C, Ayers SE, Al-Assaad Z, Carnevale K, Berger FG, Peña MM HW. Period 2 mutation accelerates ApcMin/+ tumorigenesis. Mol cancer Res. 2008;6(11):1786–93.

55. Vogelstein B. Genetic alterations during colorectal-tumor development. NEJM. 1988;319(9):525–32.

56. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, Silliman N, Szabo S, Dezso Z, Ustyanksky V, Nikolskaya T, Nikolsky Y, Karchin R, Wilson P a, Kaminker JS, Zhang Z, Croshaw R, Willis J, Dawson D, Shipitsin M, Willson JK V, Sukumar S, Polyak K, Park BH, Pethiyagoda CL, Pant PVK, Ballinger DG, Sparks AB, Hartigan J, Smith DR, Suh E, Papadopoulos N, Buckhaults P, Markowitz SD, Parmigiani G, Kinzler KW, Velculescu VE, Vogelstein B. The genomic landscapes of human breast and colorectal cancers. Science. 2007 Nov 16;318(5853):1108–13.

57. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology. 2010 Jun;138(6):2059–72.

58. Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol. 2011 Jan;6:479–507.

59. Neil J. Ganem1, Susana A. Godinho1 and DP. A Mechanism Linking Extra Centrosomes to Chromosomal Instability. Nature. 2010;460(7252):278–82.

60. Plentz RR, Wiemann SU, Flemming P, Meier PN, Kubicka S, Kreipe H, Manns MP, Rudolph KL. Telomere shortening of epithelial cells characterises the adenoma - carcinoma transition of human colorectal cancer. 2003;1304–8.

61. Katayama S, Shiota G, Oshimura M et a. Clinical usefulness of telomerase activity and telomere length in the preoperative diagnosis of gastric and colorectal cancer. Cancer Res Clin Oncol. 1999;(April 1996):405–10.

62. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-bigas MA, Fodde R, Ranzani GN. A National Cancer Institute Workshop on Microsatellite Instability for Cancer Detection and Familial Predisposition : Development of International Criteria for the Determination of Microsatellite Instability in Colorectal Cancer A National Cancer Institut. Cancer Res. 1998;(58):5248–57.

63. Kurzawski G, Suchy J, Debniak T, Kładny J, Lubiński J. Importance of microsatellite instability (MSI) in colorectal cancer: MSI as a diagnostic tool. Ann Oncol. 2004 Jan;15 Suppl 4(Supplement 4):iv283-4.

64. Kakar S, Burgart LJ, Thibodeau SN, Rabe KG, Petersen GM, Goldberg RM, Lindor NM. Frequency of loss of hMLH1 expression in colorectal carcinoma increases with advancing age. Cancer. 2003 Mar 15;97(6):1421–7.

65. Miyaki M, Iijima T, Kimura J, Yasuno M, Mori T, Hayashi Y, Koike M. Frequent Mutation of β -Catenin and APC Genes in Primary Colorectal Tumors from Patients with Hereditary Nonpolyposis Colorectal Cancer Advances in Brief Frequent Mutation of [illegible] -Catenin and APC Genes in Primary Colorectal Tumors from Patients with Heredit. cancer res. 1999;(59):4506–9.

66. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003 Mar;33 Suppl(march):245–54.

67. Issa J. CpG island methylator phenotype in cancer. Nat Rev Cancer. 2004;4(December):988–93.

68. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med. 2003 Nov 20;349(21):2042–54.

69. Shen L, Ahuja N, Shen Y, Habib NA, Toyota M, Rashid A, Issa JJ. DNA Methylation and Environmental Exposures in Human Hepatocellular Carcinoma. J Natl Cancer Inst. 2002;77030(4):755–61.

70. Toyota M, Ohe-toyota M, Ahuja N, Issa JJ. Distinct genetic profiles in colorectal tumors with or without the CpG island methylator phenotype. 1999;

71. Aoki K, Taketo MM. Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene. J Cell Sci. 2007 Oct 1;120(Pt 19):3327–35.

72. Polakis P. Wnt signaling in cancer. Cold Spring Harb Perspect Biol. 2012 May;4(5).

73. Santini D, Loupakis F, Vincenzi B, Floriani I, Stasi I, Canestrari E, Rulli E, Maltese PE, Andreoni F, Masi G, Graziano F, Baldi GG, Salvatore L, Russo A, Perrone G, Tommasino MR, Magnani M, Falcone A, Tonini G, Ruzzo A. High concordance of KRAS status between primary colorectal tumors and related metastatic sites: implications for clinical practice. Oncologist. 2008 Dec;13(12):1270–5.

74. Chan TL, Zhao W, Project CG. BRAF and KRAS Mutations in Colorectal Hyperplastic Polyps and Serrated Adenomas BRAF and KRAS Mutations in Colorectal Hyperplastic Polyps and. cancer Res. 2003;(63):4878–81.

75. Imamura Y, Morikawa T, Liao X, Lochhead P, Kuchiba A, Yamauchi M, Qian ZR, Nishihara R, Meyerhardt J a, Haigis KM, Fuchs CS, Ogino S. Specific mutations in KRAS codons 12 and 13, and patient prognosis in 1075 BRAF wild-type colorectal cancers. Clin Cancer Res. 2012 Sep 1;18(17):4753–63.

76. Guastadisegni C, Colafranceschi M, Ottini L, Dogliotti E. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. Eur J Cancer. 2010 Oct;46(15):2788–98.

77. Roth AD, Tejpar S, Delorenzi M, Yan P, Fiocca R, Klingbiel D, Dietrich D, Biesmans B, Bodoky G,

Barone C, Aranda E, Nordlinger B, Cisar L, Labianca R, Cunningham D, Van Cutsem E, Bosman F. Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. J Clin Oncol. 2010 Jan 20;28(3):466–74.

78. Andreyev HJN, Norman AR, Cunningham D, Oates J, Dix BR, Iacopetta BJ, Young J, Walsh T, Ward R. Kirsten ras mutations in patients with colorectal cancer : the " RASCAL II " study. 2001;85(May):692–6.

79. Tan C, Du X. KRAS mutation testing in metastatic colorectal cancer. World J Gastroenterol. 2012 Oct 7;18(37):5171–80.

80. Conlin a, Smith G, Carey F a, Wolf CR, Steele RJC. The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma. Gut. 2005 Sep;54(9):1283–6.

81. Markowitz SD, Bertagnolli MM. Molecular Basis of Colorectal Cancer. NEJM. 2009;361:2449–60.

82. R. J. C. STEELE, A. M. THOMPSON PAH and DPL. The p53 tumour suppressor gene. Br J Surg. 1998;85:1460–7.

83. Baker SJ, Preisinger AC, Jessup JM, Paraskeva C, Markowitz S, Willson JK V, Hamilton S, Vogelstein B. p53 Gene Mutations Occur in Combination with 17p Allelic Deletions as Late Events in Colorectal Tumorigenesis Advances in Brief p53 Gene Mutations Occur in Combination with 17p Allelic Deletions as Late Events in Colorectal Tumorigenesis1. Cancer Res. 1990;50:7717–22.

84. Cha YI, DuBois RN. NSAIDs and cancer prevention: targets downstream of COX-2. Annu Rev Med. 2007 Jan;58:239–52.

85. Sandler R, Halabi S. A randomized trial of aspirin to prevent colorectal adenomas in patients with previous colorectal cancer. NEJM 2003;248:883–90.

86. Steinbach G, Lynch P. The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. NEJM 2000;342:1946–52.

87. Wilson JMG JG. Principles and Practice of Screening for Disease. Public Health Papers 34. http://www.who.int/iris/handle/10665/37650

88. Pawa N, Arulampalam T, Norton JD. Screening for colorectal cancer: established and emerging modalities. Nat Rev Gastroenterol Hepatol. 2011 Dec;8(12):711–22.

89. Tappenden P, Chilcott J, Eggington S, Patnick J, Sakai H, Karnon J. Option appraisal of population-based colorectal cancer screening programmes in England. Gut. 2007 May;56(5):677–84.

90. Altobelli E, Lattanzi a, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. Prev Med (Baltim). 2014 May;62:132–41.

91. Segnan, N., Patnick, J., Von Karsa, L. European guidelines for quality assurance in colorectal cancer screening and diagnosis: Overview and introduction to the full Supplement publication. Endoscopy. 2013; 45(1): 51–59.

92. Juillerat P, Peytremann-Bridevaux I, Vader J-P, Arditi C, Schusselé Filliettaz S, Dubois RW,

Gonvers J-J, Froehlich F, Burnand B, Pittet V. Appropriateness of colonoscopy in Europe (EPAGE II). Presentation of methodology, general results, and analysis of complications. Endoscopy. 2009 Mar;41(3):240–6.

93. Sawbridge D, Probert C. Population-based screening in colorectal cancer - current practice and future developments: faecal biomarkers review. J Gastrointestin Liver Dis. 2014 Jun;23(2):195–202.

94. Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull B a, Ross ME. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med. 2004 Dec 23;351(26):2704–14.

95. Collins JF, Lieberman DA, Durbin TE, Weiss DG, Cooperative A. Annals of Internal Medicine Article Accuracy of Screening for Fecal Occult Blood on a Single Stool Sample Obtained by Digital Rectal Examination : A Comparison with. Ann Intern Med. 2005 Jan 18;142(2):81-5..

96. Lieberman DA WDVACSG 380. One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon. NEJM. 2001;345(8):555–60.

97. Guittet L, Bouvier V, Mariotte N, Vallee JP, Arsène D, Boutreux S, Tichet J, Launoy G. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. Gut. 2007 Feb;56(2):210–4.

98. Hol L, Wilschut J a, van Ballegooijen M, van Vuuren a J, van der Valk H, Reijerink JCIY, van der Togt a CM, Kuipers EJ, Habbema JDF, van Leerdam ME. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer. 2009 Apr 7;100(7):1103–10.

99. Kaul a, Shah a, Magill FH, Hawkins S a, Skaife P. Immunological faecal occult blood testing: a discriminatory test to identify colorectal cancer in symptomatic patients. Int J Surg. 2013 Jan;11(4):329–31.

100. Nagasaka T, Tanaka N, Cullings HM, Sun D-S, Sasamoto H, Uchida T, Koi M, Nishida N, Naomoto Y, Boland CR, Matsubara N, Goel A. Analysis of fecal DNA methylation to detect gastrointestinal neoplasia. J Natl Cancer Inst. 2009 Sep 16;101(18):1244–58.

101. Loktionov a, O'Neill IK, Silvester KR, Cummings JH, Middleton SJ, Miller R. Quantitation of DNA from exfoliated colonocytes isolated from human stool surface as a novel noninvasive screening test for colorectal cancer. Clin Cancer Res. 1998 Feb;4(2):337–42.

102. Subramanian S, Klosterman M, Amonkar MM, Hunt TL. Adherence with colorectal cancer screening guidelines: a review. Prev Med (Baltim). 2004 May;38(5):536–50.

103. von Wagner C, Good a, Wright D, Rachet B, Obichere a, Bloom S, Wardle J. Inequalities in colorectal cancer screening participation in the first round of the national screening programme in England. Br J Cancer. 2009 Dec 3;101 Suppl(S2):S60-3.

104. Gimeno García AZ: Factors influencing colorectal cancer screening participation. Gastroenterol Res Pr. 2012; Article ID 483417.

105. Adler A, Geiger S, Keil A, Bias H, Schatz P, deVos T, Dhein J, Zimmermann M, Tauber R, Wiedenmann B. Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany. BMC Gastroenterol. 2014 Jan;14(1):183.

106. Robb K, Stubbings S, Ramirez A, Macleod U, Austoker J, Waller J, Hiom S WJ. Public awareness of cancer in Britain: a population-based survey of adults. Br J Cancer. 2009;101:18–23.

107. McCaffery K, Wardle J WJ. Knowledge, attitudes, and behavioral intentions in relation to the early detection of colorectal cancer in the United Kingdom. Prev Med (Baltim). 2003;36:525–35.

108. Berkowitz Z, Hawkins NA, Peipins LA, White MC NM. Beliefs, risk perceptions, and gaps in knowledge as barriers to colorectal cancer screening in older adults. J Am Geriatr Assoc. 2008;56:307–14.

109. Stacy R, Torrence WA MC. Perceptions of knowledge, beliefs, and barriers to colorectal cancer screening. J cancer Educ. 2008;23:238–40.

110. Gimeno-García AZ, Quintero E, Nicolás-Pérez D, Parra-Blanco A J-SA. Impact of an educational video-based strategy on the behavior process associated with colorectal cancer screening: a randomized controlled study. Cancer Epidemiol. 2009;33:216–22.

111. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal. 1998;317(August):559–65.

112. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am J Gastroenterol. 2008 Jun;103(6):1541–9.

113. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal Cancer Mortality : Effectiveness of Biennial. J Natl Cancer Inst. 1999;91(5):3–6.

114. Mandel, J.S., Church, T.R., Bond, J.H. et al. T H E E F F EC T OF F ECAL OCCULT- BL OOD SC R EENING ON TH E INCID ENC E OF COLOR EC TA L CA NCER THE EFFECT OF FECAL OCCULT-BLOOD SCREENING ON THE INCIDENCE OF COLORECTAL CANCER. NEJM. 2000;1603–7.

115. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, Fedato C, Saugo M, Dei Tos a. P. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut. 2014;64(5):784–90.

116. Kronborg O, Jørgensen OD, Fenger C, Rasmussen M. Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds. Scand J Gastroenterol. 2004 Sep;39(9):846–51.

117. McClements PL, Madurasinghe V, Thomson CS, Fraser CG, Carey F a, Steele RJC, Lawrence G, Brewster DH. Impact of the UK colorectal cancer screening pilot studies on incidence, stage distribution and mortality trends. Cancer Epidemiol. 2012 Aug;36(4):e232-42.

118. Løberg M, Kalager M, Holme O, Hoff G, Adami H-O, Bretthauer M. Long-Term Colorectal-Cancer Mortality after Adenoma Removal. N Engl J Med. 2014 Aug 28;371(9):799–807.

119. Morris EJ a, Whitehouse LE, Farrell T, Nickerson C, Thomas JD, Quirke P, Rutter MD, Rees C, Finan PJ, Wilkinson JR, Patnick J. A retrospective observational study examining the characteristics and outcomes of tumours diagnosed within and without of the English NHS Bowel Cancer Screening Programme. Br J Cancer. 2012 Aug 21;107(5):757–64.

120. Nnoaham KE, Lines C. Modelling future capacity needs and spending on colonoscopy in the English bowel cancer screening programme. Gut. 2008 Sep;57(9):1238–45.

121. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, Parkin DM, Wardle J, Duffy SW, Cuzick J. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375:1624–33.

122. Ferlay J, Shin HR BF et al. GLOBOCAN 2008 v2.0. Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 10 2010. 2008.

123. El-Serag HB. Hepatocellular Carcinoma. NEJM. 2011;365:1118–27.

124. Bosetti C, Turati F, La Vecchia C. Hepatocellular carcinoma epidemiology. Best Pract Res Clin Gastroenterol. 2014 Aug;28(5):753–70.

125. De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D et al. Cancer survival in Europe 1999e2007 by country and age: results of EUROCAREe5-a population-based study. Lancet Oncol. 2014;15:23–34.

126. Chen C-J, Yang H-I, Iloeje UH. Hepatitis B virus DNA levels and outcomes in chronic hepatitis B. Hepatology. 2009 May;49(5 Suppl):S72-84.

127. Ryder SD. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults. Gut. 2003 May 1;52(90003):1iii–8.

128. Laupacis A, Feeny D, Detsky AS, Tugweli PX. How attractive does a new technology have to be to warrant adoption and utilization ? Tentative guidelines for using clinical and economic evaluations. CAN MED ASSOC J. 1992;146(4):473–81.

129. Dufour JF, Greten TF, Raymond E, Roskams T, De T, Ducreux M, Mazzaferro V, Governing E. Clinical Practice Guidelines EASL – EORTC Clinical Practice Guidelines : Management of hepatocellular carcinoma European Organisation for Research and Treatment of Cancer. Vol. 56. 2012.

130. Sherman M. Surveillance for hepatocellular carcinoma. Best Pract Res Clin Gastroenterol. 2014 Oct;28(5):783–93.

131. Chen JG, Parkin DM, Chen QG, Lu JH, Shen QJ, Zhang BC. Screening for liver cancer: results of a randomised controlled trial in Qidong, China. J Medcial Screen. 2003;10:204–9.

132. Zhang BH, Yang BH TZ. Randomized controlled trial of screening for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2004;130:417–22.

133. Poustchi H, Farrell GC, Strasser SI, Lee AU, McCaughan GW GJ. Feasibility of conducting a randomized control trial for liver cancer screening: is a randomized controlled trial for liver cancer screening feasible or still needed? Hepatology. 2001;54:1998–2004.

134. Yeh YP, Hu TH, Cho PY, Chen HH, Yen AM, Chen SL et al. Evaluation of abdominal ultrasonography mass screening for hepatocellular carcinoma in Taiwan. Hepatology. 2014;59:1840–9.

135. Chan AC, Poon RT, Ng KK, Lo CM, Fan ST WJ. Changing paradigm in the management of hepatocellular carcinoma improves the survival benefit of early detection by screening. Ann Surg. 2008;247:666–73.

136. Wong GL, Wong VW, Tan GM, Ip KI, Lai WK, Li YW et al. Surveillance programme for hepatocellular carcinoma improves the survival of patients with chronic viral hepatitis. Liver Int. 2008;28:999–1003.

137. Nouso K, Tanaka H, Uematsu S, Shiraga K, Okamoto R, Onishi H et al. Cost-effectiveness of the surveillance program of hepatocellular carcinoma depends on the medical circumstances. J Gastroenterol Hepatol. 2008;23:437–44.

138. Cucchetti A, Trevisani F, Cescon M, Ercolani G, Farinati F, Poggio PD et al. Cost-effectiveness of semi-annual surveillance for hepatocellular carcinoma in cirrhotic patients of the Italian Liver Cancer population. J Hepatol. 2012;56:1089–96.

139. Andersson KL, Salomon JA, Goldie SJ CR. Cost effectiveness of alternative surveillance strategies for hepatocellular carcinoma in patients with cirrhosis. Clin Gastroenterol Hepatol. 2008;6:1418–24.

140. Bolondi L, Sofia S, Siringo S, Gaiani S, Casali A, Zironi G, Piscaglia F, Gramantieri L, Zanetti M, Sherman M, Bologna U. Surveillance programme of cirrhotic patients for early diagnosis and treatment of hepatocellular carcinoma : a cost e V ectiveness analysis. Gut. 2001;48:251–9.

141. Bolondi L. Screening for hepatocellular carcinoma in cirrhosis. J Hepatol. 2003 Dec;39(6):1076–84.

142. Xu J, Liu C, Zhou L, Tian F, Tai M-H, Wei J-C, Qu K, Meng F-D, Zhang L-Q, Wang Z-X, Zhang J-Y, Chang H-L, Liu S-N, Xu X-S, Song Y-Z, Liu J, Zhang P. Distinctions Between Clinicopathological Factors and Prognosis of Alpha-fetoprotein Negative and Positive Hepatocelluar Carcinoma Patients. Asian Pacific J Cancer Prev. 2012 Feb 29;13(2):559–62.

143. Abbasi A, Bhutto AR, Butt N, Munir SM. Original Article Corelation of serum alpha fetoprotein and tumor size in hepatocellular carcinoma. J Pakistan Med Assoc. 2012;62(1):33–6.

144. Singal A, Volk ML, Waljee A, Salgia R, Higgins P, Rogers MA et al. Metaanalysis: surveillance with ultrasound for early-stage hepatocellular carcinoma in patients with cirrhosis. Aliment Pharmacol Ther. 2009;30:37–47.

145. Sato T, Tateishi R, Yoshida H, Ohki T, Masuzaki R, Imamura J, Goto T, Kanai F, Obi S, Kato N, Shiina S, Kawabe T, Omata M. Ultrasound surveillance for early detection of hepatocellular carcinoma among patients with chronic hepatitis C. Hepatol Int. 2009 Dec;3(4):544–50.

146. Lencioni R, Piscaglia F BL. Contrast-enhanced ultrasound in the diagnosis of hepatocellular carcinoma. J Hepatol. 2008;48:848–57.

147. Han KH, Kim Do Y, Park JY, Ahn SH, Kim J, Kim SU et al. Survival of hepatocellular carcinoma patients may be improved in surveillance interval not more than 6 months compared with more than 6 months: a 15-year prospective study. J Clin Gastroenterol. 2013;47:538–44.

148. Santi V, Trevisani F, Gramenzi A, Grignaschi A, Mirici-Cappa F, Del Poggio P et al. Semiannual surveillance is superior to annual surveillance for the detection of earlyhepatocellular carcinoma and patient survival. J Hepatol. 2010;53:291–7.

149. Trinchet JC, Chaffaut C, Bourcier V, Degos F, Henrion J, Fontaine H et al. Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3- and 6-month periodicities. Hepatology. 2011;54:1987–97.

150. Kansagara D, Papak J, Pasha AS, O'Neil M, Freeman M, Relevo R et al. Screening for hepatocellular carcinoma in chronic liver disease: a systematic review. Ann Intern Med. 2014;161:261–9.

151. Singal AG, Pillai A TJ. Early detection, curative treatment, and survival rates for hepatocellular

carcinoma surveillance in patients with cirrhosis: a meta-analysis. PloS Med. 2014;11:e1001624.

152. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. Lancet. 2012;379(9822):1245–55.

153. Grisham JW. Molecular genetic alterations in primary hepatocellular neoplasms: hepatocellular adenoma, hepatocellular carcinoma, and hepatoblastoma. Mol Basis Hum Cancer. 2001;269–346.

154. Bosch, F.X., Ribes, J. & Borràs J. Epidemiology of primary liver cancer. Semin Liver Dis. 1999;19:271–85.

155. Moradpour D, Blum HE. Pathogenesis of hepatocellular carcinoma. Eur J Gastroenterol Hepatol. 2005 May;17(5):477–83.

156. Thorgeirsson S, Grisham J. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet. 2002;31(august):339–46.

157. Nault J-C. Pathogenesis of hepatocellular carcinoma according to aetiology. Best Pract Res Clin Gastroenterol. 2014 Oct;28(5):937–47.

158. Beasley RP, Hwang L-Y, Lin C-C CC-S. Hepatocellular carcinoma and hepatitis B virus – a prospective study of 22707 men in Taiwan. Lancet. 1981;ii:1129–33.

159. Neuveut, C. Wei, Y. Buendia M. Mechanisms of HBV-related hepatocarcinogenesis. J Hepatol. 2010;52:594–604.

160. Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet. 2012;44:765–9.

161. Brechot C. Pathogenesis of hepatitis B virus-related hepatocellular carcinoma: old and new paradigms. Gastroenterology. 2004;127:S56-61.

162. Dejean A, Bougueleret L, Grzeschik KH TP. Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. Nature. 1986;322:70–2.

163. Wang J, Chenivesse X, Henglein B B chot C. Hepatitis B virus i ntegration in a cyclin A gene in a hepatocellular carcinoma. Nature. 1990;343:555–7.

164. Seifer M, Hohne M, Schaefer S GW. In vitro tumorigenicity of hepatitis B virus DNA and HBx protein. J Hepatol. 1991;13:S61-5.

165. Benhenda S, Cougot D, Buendia MA NC. Hepatitis B virus X protein molecular functions and its role in virus life cycle and pathogenesis. Adv Cancer Res. 2009;103:75–109.

166. Wen Y, Golubkov VS, Strongin AY, Jiang W RJ. Interaction of hepatitis B viral oncoprotein with cellular target HBXIP dysregulates centrosome dynamics and mitotic spindle formation. J Biol Chem. 2008;283:2793–803.

167. Ray RB, Steele R, Meyer K RR. Transcriptional repression of p53 promoter by hepatitis C virus core protein. J Biol Chem. 1997;272:10983–6.

168. Joo M, Hahn YS, Kwon M, Sadikot RT, Blackwell TS CJ. Hepatitis C virus core protein suppresses NF-kappaB activation and cyclooxygenase-2 expression by direct interaction with IkappaB kinase beta. J Virol. 2005;79:7648–57.

169. Morgan TR, Mandayam S JM. Alcohol and hepatocellular carcinoma. Gastroenterology. 2004;127:S87-96.

170. De Bac C, Stroffolini T, Gaeta GB, Taliani G GG. Pathogenic factors in cirrhosis with and without hepatocellular carcinoma: a multicenter Italian study. Hepatology. 1994;20:1225–30.

171. Tsukuma H, Hiyama T, Tanaka S, Nakao M, Yabuuchi T, Kitamura T et al. Risk factors for hepatocellular carcinoma among patients with chronic liver disease. NEJM. 1993;328:1797–801.

172. Stickel F, Schuppan D, Hahn EG SH. Cocarcinogenic effects of alcohol in hepatocarcinogenesis. Gut. 2002;51:13–9.

173. Seitz HK SF. Molecular mechanisms of alcohol-mediated carcinogenesis. Nat Rev Cancer. 2007;7:599–612.

174. Bradford BU, Kono H, Isayama F, Kosyk O, Wheeler MD, Akiyama TE et al. Cytochrome P450 CYP2E1, but not nicotinamide adenine dinucleotide phosphate oxidase, is required for ethanol-induced oxidative DNA damage in rodent liver. Hepatology. 2005;41:336–44.

175. Mato JM, Martínez-Chantar ML, Lu SC. Methionine metabolism and liver disease. Annu Rev Nutr. 2008 Jan;28:273–93.

176. Avila MA, Berasain C, Torres L, Martin-duce A, Corrales FJ, Yang H, Prieto J, Lu SC, Caballeria J, Rodcs J, Mato JM. Reduced mRNA abundance of the main enzymes involved in methionine metabolism in human liver cirrhosis and hepatocellular carcinoma. J Hepatol. 2000;(4):907–14.

177. Lu SC, Martínez-Chantar ML, Mato JM. Methionine adenosyltransferase and S-adenosylmethionine in alcoholic liver disease. J Gastroenterol Hepatol. 2006 Oct;21 Suppl 3:S61-4.

178. Ertle J, Dechene A, Sowa JP, Penndorf V, Herzer K, Kaiser G et al. Non-alcoholic fatty liver disease progresses to hepatocellular carcinoma in the absence of apparent cirrhosis. Int J cancer. 2011;128:2436–43.

179. Paradis V, Zalinski S, Chelbi E, Guedj N, Degos F, Vilgrain V et al. Hepatocellular carcinomas in patients with metabolic syndrome often develop without significant liver fibrosis: a pathological analysis. Hepatology. 2009;49:851–9.

180. Siddique A KK. Insulin resistance and other metabolic risk factors in the pathogenesis of hepatocellular carcinoma. Clin liver Dis. 2011;15:281–96.

181. Kamada Y, Matsumoto H, Tamura S, Fukushima J, Kiso S, Fukui K et al. Hypoadiponectinemia accelerates hepatic tumor formation in a nonalcoholic steatohepatitis mouse model. J Hepatol. 2007;47:556–64.

182. Gallagher EJ LD. Minireview: IGF, insulin, and cancer. Endocrinology. 2011;152:2546–51.

183. Nishida N, Fukuda Y, Komeda T, Kita R, Sando T, Furukawa M et al. Amplification and overexpression of the cyclin D1 gene in aggressive human hepatocellular carcinoma. Cancer Res. 1994;54:3107–10.

184. Woo HG, Wang XW, Budhu A, Kim YH, Kwon SM, Tang ZY et al. Association of TP53 mutations with stem cell-like gene expression and survival of patients with hepatocellular carcinoma.

Gastroenterology. 2011;140:1063–70.

185. Bressac B, Kew M, Wands J OM. Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa. Nature. 1991;350:429–31.

186. Ozturk M. Genetic aspects of hepatocellular carcinogenesis. Semin Liver Dis. 1999;19:235–42.

187. Farazi PA, Glickman J, Jiang S, Yu A, Rudolph KL DR. Differential impact of telomere dysfunction on initiation and progression of hepatocellular carcinoma. Cancer Res. 2003;63:5021–7.

188. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89–95.

189. WHO. WHO International Programme on Chemical Safety Biomarkers in Risk Assessment: Validity and Validation. 2001. http://www.who.int/iris/handle/10665/42363

190. Newton KF, Newman W, Hill J. Review of biomarkers in colorectal cancer. Colorectal Dis. 2012 Jan;14(1):3–17.

191. Tanaka T, Tanaka M, Tanaka T, Ishigamori R. Biomarkers for colorectal cancer. Int J Mol Sci. 2010 Jan;11(9):3209–25.

192. Shah R, Jones E, Vidart V, Kuppen PJK, Conti J a, Francis NK. Biomarkers for early detection of colorectal cancer and polyps: systematic review. Cancer Epidemiol Biomarkers Prev. 2014 Sep;23(9):1712–28.

193. Boynton, K.A.; Summerhayes, I.C.; Ahlquist, D.A.; Shuber AP. DNA integrity as a potential marker for stool-based detection of colorectal cancer. Clin Chem. 2003;49:1058–1065.

194. Tonus C, Sellinger M, Koss K, Neupert G. Faecal pyruvate kinase isoenzyme type M2 for colorectal cancer screening: a meta-analysis. World J Gastroenterol. 2012 Aug 14;18(30):4004–11.

195. Mazurek S. Pyruvate kinase type M2: a key regulator of the metabolic budget system in tumor cells. Int J Biochem Cell Biol. 2011;43(7):969–80.

196. Koss K, Maxton D JJ. Fecal dimeric M2 pyruvate kinase in colorectal cancer and polyps correlates with tumor staging and surgical intervention. Colorectal Dis. 2008;10:244–8.

197. Liu Z, Zhang Y, Niu Y, Li K, Liu X, Chen H, Gao C. A systematic review and meta-analysis of diagnostic and prognostic serum biomarkers of colorectal cancer. PLoS One. 2014 Jan;9(8):e103910.

198. Fletcher RH. Carcinoembryonic antigen. Ann Intern Med. 1986;104:66–73.

199. Duffy MJ. CEA as a marker for colorectal cancer: is it clinically useful? Clin Chem. 2001;47:624–630.

200. Goldstein MJ MM. Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer. Cancer Investig. 2005;23:338–351.

201. Prager GW, Braemswig KH, Martel A, Unseld M, Heinze G, Brodowicz T, Scheithauer W, Kornek G ZC. Baseline carcinoembryonic antigen (CEA) serum levels predict bevacizumab-based treatment response in metastatic colorectal cancer. Cancer Sci. 2014;105(8):996–1001.

202. Hara M, Sato M, Takahashi H, Takayama S, Okada Y, Nagasaki T TH. Carcinoembryonic antigen elevation in post-hepatectomy patients with colorectal cancer liver metastasis indicates recurrence with high accuracy. Hepatogastroenterology. 2013;60(128):1935–9.

203. Carpelan-Holmstro¨m M, Louhimo J, Stenman UH, Alfthan H, Ja¨rvinen H E, Al. CEA, CA 242, CA-19-9, CA 72-4 and hCGbin the diagnosis of recurrent colorectal cancer. Tumour Biol. 2004;25:228–34.

204. Sorensen, N.M.; Schrohl, A.S.; Jensen, V.; Christensen, I.J.; Nielsen, H.J.; Brunner N. Comparative studies of tissue inhibitor of metalloproteinases-1 in plasma, serum and tumour tissue extracts from patients with primary colorectal cancer. Scand J Gastroenterol. 2008;43:186–191.

205. Holten-Andersen, M.N.; Fenger, C.; Nielsen, H.J.; Rasmussen, A.S.; Christensen, I.J.; Brunner, N.; Kronborg O. Plasma TIMP-1 in patients with colorectal adenomas: A prospective study. Eur J Cancer. 2004;40:2159–2164.

206. Holten-andersen MN, Stephens RW, Nielsen HJ, Murphy G, Christensen IJ, Stetler-stevenson W, Bru N. High Preoperative Plasma Tissue Inhibitor of Metalloproteinase-1 Levels Are Associated with Short Survival of Patients with. Clin Cancer Res. 2000;6(November):4292–9.

207. Kim H-J, Yu M-H, Kim H-G, Byun J-H, Lee C. Noninvasive molecular biomarkers for the detection of colorectal cancer. BMB Rep. 2008 Oct 31;41(10):685–92.

208. Wu, C. C., Chen, H. C., Chen, S. J., Liu, H. P., Hsieh Y, Y., Yu, C. J., Tang, R., Hsieh, L. L., Yu, J. S. and Chang Y, S. Identification of collapsin response mediator protein-2 as a potential marker of colorectal carcinoma by comparative analysis of cancer cell secretomes. Proteomics. 2008;8:316–32.

209. Mishra PJ. MicroRNAs as promising biomarkers in cancer diagnostics. Biomark Res. 2014 Jan;2(1):19.

210. Tazawa, H.; Tsuchiya, N.; Izumiya, M.; Nakagama H. Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. Proc Natl Acad Sci U S A. 2007;104:15472–15477.

211. Ju J, Jiang J FA. miRNA: the new frontier in cancer medicine. Futur Med Chem. 2013;5(9):983–985.

212. Song B, Wang Y, Titmus MA, Botchkina G, Formentini A, Kornmann M JJ. Molecular mechanism of chemoresistance by miR-215 in osteosarcoma and colon cancer cells. Mol cancer Res. 2010;9:96.

213. Schepeler, T.; Reinert, J.T.; Ostenfeld, M.S.; Christensen, L.L.; Silahtaroglu, A.N.; Dyrskjot, L.; Wiuf, C.; Sorensen, F.J.; Kruhoffer, M.; Laurberg, S.; Kauppinen, S.; Orntoft, T.F.; Andersen CL. Diagnostic and prognostic microRNAs in stage II colon cancer. Cancer Res. 2008;68:6416–6424.

214. Diaz, R.; Silva, J.; Garcia, J.M.; Lorenzo, Y.; Garcia, V.; Pena, C.; Rodriguez, R.; Munoz, C.; Garcia, F.; Bonilla, F.; Dominguez G. Deregulated expression of miR-106a predicts survival in human colon cancer patients. Genes, Chromosom Cancer. 2008;47:794–802.

215. Siena S, Sartore-Bianchi A, Di Nicolantonio F BJ, A. B. Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. J Natl cancer institute. 2009;101:1308–24.

216. Van Cutsem E, Peeters M SS et al. Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. J Clin Oncol. 2007;25:1658–64.

217. Bokemeyer C, Bondarenko I MA et al. Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer. J Clin Oncol. 2009;27:663–71.

218. Di Nicolantonio F, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S BA. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. J Clin Oncol. 2008;10:5705–12.

219. Gupta S, Bent S KJ. Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. Ann Intern Med. 2003;139:46–50.

220. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P et al. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. J Hepatol. 2001;34:570–5.

221. Koike Y, Shiratori Y, Sato S, Obi S, Teratani T, Imamura M, Yoshida H, Shiina S, Omata M. Des-gamma -Carboxy Prothrombin As a Useful Predisposing Factor for the Development of Portal Venous Invasion in Patients with Hepatocellular Carcinoma A Prospective Analysis of 227 Patients. Cancer. 2001;91:561–569.

222. Rich N, Singal AG. Hepatocellular carcinoma tumour markers: current role and expectations. Best Pract Res Clin Gastroenterol. 2014 Oct;28(5):843–53.

223. Richardson P, Duan Z, Kramer J, Davila JA, Tyson GL, El-Serag HB et al. Determinants of serum alpha-fetoprotein levels in hepatitis C-infected patients. Clin Gastroenterol Hepatol. 2012;10:428–33.

224. Gopal P, Yopp AC,Waljee AK, Chiang J, Nehra M, Kandunoori P et al. Factors that affect accuracy of alpha-fetoprotein test in detection of hepatocellular carcinoma in patients with cirrhosis. Clin Gastroenterol Hepatol. 2013;

225. Shiraki K, Takase K, Tameda Y, Hamada M, Kosaka Y, Nakano T et al. A clinical study of lectin-reactive alpha-fetoprotein as an early indicator of hepatocellular carcinoma in the follow-up of cirrhotic patients. Hepatology. 1995;22:802–7.

226. Sterling RK, Jeffers L, Gordon F, Sherman M, Venook AP, Reddy KR et al. Clinical utility of AFP-L3% measurement in North American patients with HCV-related cirrhosis. Am J Gastroenterol. 2007;102:2196–205.

227. Marrero JA, Feng Z,Wang Y, Nguyen MH, Befeler AS, Roberts LR et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. Gastroenterology. 2009;137:110–8.

228. Oda K, Ido A, Tamai T, Matsushita M, Kumagai K, Mawatari S et al. Highly sensitive lens culinaris agglutinin-reactive alphafetoprotein is useful for early detection of hepatocellular carcinoma in patients with chronic liver disease. Oncol Rep. 2011;26:1227–33.

229. Behne T, Copur MS. Biomarkers for hepatocellular carcinoma. Int J Hepatol. 2012 Jan;2012:859076.

230. Weitz, I.C.a, Liebman H. Des-γ-carboxy (abnormal) prothrombin and hepatocellular carcinoma: A critical review. Hepatology. 1993;18:990–7.

231. Marrero JA, Su GL, Wei W, Emick D, Conjeevaram HS, Fontana RJ et al. Des-gamma carboxyprothrombin can differentiate hepatocellular carcinoma from nonmalignant chronic liver disease in american patients. Hepatology. 2003;37:1114–21.

232. Izuno K, Fujiyama S, Yamasaki K, Sato M ST. Early detection of hepatocellular carcinoma associated with cirrhosis by combined assay of des-gamma-carboxy prothrombin and alpha-fetoprotein: a prospective study. Hepatogastroenterology. 1995;24:387–93.

233. Ishii M, Gama H, Chida N, Ueno Y, Shinzawa H, Takagi T et al. Simultaneous measurements of serum alpha-fetoprotein and protein induced by vitamin K absence for detecting hepatocellular carcinoma. South Tohoku District Study Group. Am J Gastroenterol. 200AD;95:1036–40.

234. Lok AS, Sterling RK, Everhart JE, Wright EC, Hoefs JC, Di Bisceglie AM et al. Des-gamma-carboxy prothrombin and alphafetoprotein as biomarkers for the early detection of hepatocellular carcinoma. Gastroenterology. 2010;138:493–502.

235. Nakatsura T, Yoshitake Y, Senju S, Monji M, Komori H, Motomura Y, Hosaka S, Beppu T, Ishiko T, Kamohara H, Ashihara H, Katagiri T, Furukawa Y, Fujiyama S, Ogawa M, Nakamura Y, Nishimura Y. Glypican-3, overexpressed specifically in human hepatocellular carcinoma, is a novel tumor marker. Biochem Biophys Res Commun. 2003 Jun;306(1):16–25.

236. Capurro M, Wanless IR, Sherman M, Deboer G, Shi W, Miyoshi E, Filmus J. Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. Gastroenterology. 2003 Jul;125(1):89–97.

237. Tangkijvanich P, Chanmee T, Komtong S, Mahachai V, Wisedopas N, Pothacharoen P et al. Diagnostic role of serum glypican- 3 in differentiating hepatocellular carcinoma from non-malignant chronic liver disease and other liver cancers. J Gastroenterol Hepatol. 2010;25:129–37.

238. Shirakawa H, Suzuki H, Shimomura M, Kojima M, Gotohda N, Takahashi S, Nakagohri T, Konishi M, Kobayashi N, Kinoshita T, Nakatsura T. Glypican-3 expression is correlated with poor prognosis in hepatocellular carcinoma. Cancer Sci. 2009 Aug;100(8):1403–7.

239. Kim J, Ki SS, Lee SD, Han CJ, Kim YC, Park SH et al. Elevated plasma osteopontin levels in patients with hepatocellular carcinoma. Gastroenterology. 2006;101:2051–9.

240. Shang S, Plymoth A, Ge S, Feng Z, Rosen HR, Sangrajrang S et al. Identification of osteopontin as a novel marker for early hepatocellular carcinoma. Hepatology. 2012;55:483–90.

241. Tangkijvanich P, Tosukhowong P, Bunyongyod P, Lertmaharit S, Hanvivatvong O, Kullavanijaya P PY. Alpha-L-fucosidase as a serum marker of hepatocellular carcinoma in Thailand. Southeast Asian J Trop Med Public Heal. 1999;30:110–114.

242. Marrero JA, Romano PR, Nikolaeva O, Steel L, Mehta A, Fimmel CJ et al. GP73, a resident golgi glycoprotein, is a novel serum marker for hepatocellular carcinoma. J Hepatol. 2005;55:483–90.

243. Riener, M. Stenner, F. Liewen H. Golgi phosphoprotein 2 (GOLPH2) expression in liver tumors and its value as a serum marker in hepatocellular carcinomas. Hepatology. 2009;49:1602–9.

244. Zhou L, Liu J, Luo F. Serum tumor markers for detection of hepatocellular carcinoma. World J Gastroenterol. 2006;12(8):1175–81.

245. Song B-C, Chung Y-H, Kim J a, Choi W-B, Suh DD, Pyo S Il, Shin JW, Lee HC, Lee YS, Suh DJ. Transforming growth factor-beta1 as a useful serologic marker of small hepatocellular carcinoma. Cancer. 2002 Jan 1;94(1):175–80.

246. Giannelli G, Marinosci F, Trerotoli P, Volpe A, Quaranta M, Dentico P et al. SCCA antigen combined with alpha-fetoprotein as serologic markers of HCC. Int J Cancer. 2005;117:506–9.

247. Giannelli G, Fransvea E, Trerotoli P, Beaugrand M, Marinosci F, Lupo L et al. Clinical validation of combined serological biomarkers for improved hepatocellular carcinoma diagnosis in 961 patients. Clin Chim Acta. 2007;383:147–52.

248. Ferracin M, Veronese A, Negrini M. Micromarkers: miRNAs in cancer diagnosis and prognosis. Expert Rev Mol Diagn. 2010 Apr;10(3):297–308.

249. Zhang Y, Takahashi S, Tasaka A, Yoshima T, Ochi H CK. Involvement of microRNA-224 in cell proliferation, migration, invasion, and anti-apoptosis in hepatocellular carcinoma. J Gastroenterol Hepatol. 2013;28:565–75.

250. Su H, Yang JR, Xu T, Huang J, Xu L, Yuan Y ZS. MicroRNA-101, down-regulated in hepatocellular carcinoma, promotes apoptosis and suppresses tumorigenicity. Cancer Res. 2009;69:1135–42.

251. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. Cell. 2012 Mar 16;148(6):1258–70.

252. Shendure J. Next-generation DNA sequencing. Nat biotechology. 2008;26:1135–1145.

253. Marchesi JR, Dutilh BE, Hall N, Peters WHM, Roelofs R, Boleij A, Tjalsma H. Towards the human colorectal cancer microbiome. PLoS One. 2011 Jan;6(5):e20447.

254. Qin J, Li R, Raes J, Arumugam M, Burgdorf S, Manichanh C, Nielsen T, Pons N, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J, Hansen T, Paslier D Le, Linneberg A, Nielsen HB, Pelletier E, Renault P, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H. A human gut microbial gene catalog established by metagenomic sequencing. Nature. 2010;464(7285):59–65.

255. Jimenez, E., Marın, M.L., Martın, R., Odriozola, J.M., Olivares, M., Xaus J, Fernandez, L., and Rodrıguez JM. Is meconium from healthy newborns actually sterile? Res Microbiol. 2008;159:187–193.

256. Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight R, Angenent, L.T., and Ley RE. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci USA. 2011;108 (Suppl:4578–4585.

257. Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo G, Fierer, N., and Knight R. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci USA. 2010;107:11971–11975.

258. Vaishampayan, P.A., Kuehl, J.V., Froula, J.L., Morgan, J.L., Ochman, H. A, Francino MP. Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. Genome Biol Evol. 2010;2:53–66.

259. Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown PO. Development of the human infant intestinal microbiota. PloS Biol. 2007;5:e177.

260. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, Harris HMB, Coakley M, Lakshminarayanan B, O'Sullivan O, Fitzgerald GF, Deane J, O'Connor M, Harnedy N, O'Connor K, O'Mahony D, van Sinderen D, Wallace M, Brennan L, Stanton C, Marchesi JR, Fitzgerald AP, Shanahan F, Hill C, Ross RP, O'Toole PW. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012 Aug 9;488(7410):178–84.

261. Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., and Gordon JI. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol. 2008;6:776–788.

262. Guigoz, Y., Dore, J., Schiffrin EJ. The inflammatory status of old age can be nurtured from the intestinal environment. Curr Opin Clin Nutr Metab Care. 2008;11:13–20.

263. Lovat LB. Age related changes in gut physiology and nutritional status. Gut. 1996;38:306–309.

264. Dethlefsen, L., Huse, S., Sogin, M.L., and Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PloS Biol. 2008;6:e280.

265. Keku TO, Dulal S, Deveaux A, Jovov B, Han X. The Gastrointestinal Microbiota and Colorectal Cancer. Am J Physiol Gastrointest Liver Physiol. 2015 Mar 1;308(5):G351-63.

266. Zackular JP, Baxter NT, Iverson KD. The Gut Microbiome Modulates Colon Tumorigenesis. Am J Clin Nutr. 2013;97:1044/1052.

267. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet. 2012 Apr;13(4):260–70.

268. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol. 2014 Oct;12(10):661-72.

269. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer R-J. Review article: the role of butyrate on colonic function. Aliment Pharmacol Ther. 2008 Jan 15;27(2):104–19.

270. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T-J, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor A a, Jobin C. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science. 2012 Oct 5;338(6103):120–3.

271. Huycke MM, Abrams V and MD. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. Carcinogenesis. 2002;23:529–36.

272. Keku TO, Dulal S, Deveaux A, Jovov B, Han X. The gastrointestinal microbiota and colorectal cancer. Am J Physiol - Gastrointest Liver Physiol. 2015;308(5):G351–63.

273. Sartor RB. Microbial influences in inflammatory bowel diseases. Gastroenterology. 2008;134:577–94.

274. Grivennikov, S. I., Greten, F. R. and Karin M. Immunity, inflammation, and cancer. Cell. 2010;140:883–899.

275. Fukata M, Chen A, Vamadevan AS, Cohen J, Breglio K, Krishnareddy S, Hsu D, Xu R, Harpaz N, Dannenberg AJ, Subbaramaiah K, Cooper HS, Itzkowitz SH AM. Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status.

Gastroenterology. 2007;133:1869–1881.

276. Guarner F. Enteric flora in health and disease. Digestion. 2006;Suppl 1:5–12.

277. Pryde SE, Duncan SH, Hold GL, Stewart CS, Ã HJF. The microbiology of butyrate formation in the human colon. FEMS Microbiol Lett. 2002;217:133–9.

278. Macfarlane GT and MS. Bacteria, colonic fermentation, and gastrointestinal health. J AOAC Int. 2012;95:50–60.

279. Csordas A. Butyrate, aspirin and colorectal cancer. Eur J cancer Prev. 1996;5:221–31.

280. Segain, J.P., delaBletiere, D.R., Boureille, A., Leray, V., Gervois N, Rosales, C., Ferrier, L., Bonnet, C., Blottiere HM and G, J.P. Butyrate inhibits in£ammatory responses through NFkappaB inhibition : implications for Crohn's disease. Gut. 2000;47:397–403.

281. Smith PM, Howitt MR, Panikov N, Michaud M, Ann C, Bohlooly-y M, Glickman JN, Garrett WS. The microbial metabolites, short chain fatty acids, regulate colonic Treg cell homeostasis. Science (80- ). 2013;341:569–573.

282. Luciano L, Groos S, Busche R et al. Massive apoptosis of colonocytes induced by butyrate deprivation overloads resident macrophages and promotes the recruitment of circulating monocytes. Cell tissue Res. 2002;309:393–407.

283. Guilloteau P, Martin L, Eeckhaut V, Ducatelle R, Zabielski R, Van Immerseel F. From the gut to the peripheral tissues: the multiple effects of butyrate. Nutr Res Rev. 2010 Dec;23(2):366–84.

284. Zeng H & Briske-Anderson M. Prolonged butyrate treatment inhibits the migration and invasion potential of HT1080 tumor cells. J Nutr. 2005;135:291–295.

285. Zgouras D, Wächtershäuser A, Frings D SJ. Butyrate impairs intestinal tumor cell-induced angiogenesis by inhibiting HIF-1alpha nuclear translocation. Biochem Biophys Res Commun. 2003;300:832–8.

286. Weir TL, Manter DK, Sheflin AM, Barnett B a, Heuberger AL, Ryan EP. Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. PLoS One. 2013 Jan;8(8):e70803.

287. Gill, C. I. R. & Rowland IR. Diet and cancer: assessing the risk. Br J Nutr. 2002;88:S73–S87.

288. Windey, K., de Preter, V. & Verbeke K. Relevance of protein fermentation to gut health. Mol Nutr food Res. 2012;56:184–196.

289. Loh YH, Jakszyn P, Luben RN, Mulligan AA, Mitrou PN KK. N‑nitroso compounds and cancer incidence: The European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk Study. Am J Nutr. 2011;93:1053–1061.

290. Pegg AE. Toxicity of polyamines and their metabolic products. Chem Res Toxicol. 2013;26:1782–1800.

291. Di Martino ML, Campilongo R, Casalino M, Micheli G, Colonna B PG. Polyamines: emerging players in bacteria–host interactions. Int J Med Microbiol. 2013;303:484–491.

292. Marquet, P., Duncan, S. H., Chassard, C. B-, Donadille, A. & Flint HJ. Lactate has the potential to promote hydrogen sulphide formation in the human colon. FEMS Microbiol Lett. 2009;299:128–134.

293. Carbonero, F., Benefiel, A. C. & Gaskins HR. Contributions of the microbial hydrogen economy to colonic homeostasis. Nat Rev Gastroenterol Hepatol. 2012;9:504–518.

294. Barrasa, J. I., Olmo, N., Lizarbe, M. A. & Turnay J. Bile acids in the colon, from healthy to cytotoxic molecules. Toxicol Vitr. 2013;27:964–977.

295. Ou, J., DeLany, J. P., Zhang, M., Sharma S&, O'Keefe SJD. Association between low colonic short-chain fatty acids and high bile acids in high colon cancer risk populations. Nutr Cancer. 2012;64:34–40.

296. Azcarate-Peril MA, Sikes M and B-BJ. The intestinal microbiota, gastrointestinal environment and colorectal cancer: a putative role for probiotics in prevention of colorectal cancer? Am J Physiol Gastrointest Liver Physiol. 2011;301:G401-424.

297. Ou J, Carbonero F ZE et al. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. Am J Clin Nutr. 2013;98:111–120.

298. Ohtani N. Microbiome and cancer. Semin Immunopathol. 2014 Nov 18;65–72.

299. Castellarin M, Warren RL, Freeman JD, Warren L, Dreolini L, Krzywinski M, Strauss J, Barnes R, Watson P, Allen-vercoe E, Moore R a, Holt R a. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. 2012;299–306.

300. Kostic AD, Gevers D PC et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome Res. 2012;22:292–29.

301. Rubinstein MR, Wang X LW et al. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating Ecadherin/ β-catenin signaling via its FadA adhesin. Cell host microbe. 2013;14:195–206.

302. Zackular JP, Rogers M a. M, Ruffin MT, Schloss PD. The Human Gut Microbiome as a Screening Tool for Colorectal Cancer. Cancer Prev Res. 2014 Aug 7;1–10.

303. Baxter NT, Zackular JP, Chen GY and SP. Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. Microbiome. Microbiome. 2014 Jun 17;2:20..

304. Makivuokko H, Tiihonen K, Tynkkynen S et al. The effect of age and non-steroidal anti-inflammatory drugs on human intestinal microbiota composition. Br J Nutr. 2010;103:227–234.

305. Enck P, Zimmermann K, Rusch K et al. The effects of ageing on the colonic bacterial microflora in adults. Z Gastroenterol. 2009;47:653–658.

306. Probert CSJ, Ahmed I, Khalid T, Johnson E, Smith S, Ratcliffe N. Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases. J Gastrointestin Liver Dis. 2009 Sep;18(3):337–43.

307. Reade S MA. Optimisation of Sample Preparation for Direct SPME-GC-MS Analysis of Murine and Human Faecal Volatile Organic Compounds for Metabolomic Studies. J Anal Bioanal Tech. 2014;5(2).

308. Sonoda H, Kohnoe S, Yamazato T, Satoh Y, Morizono G, Shikata K, Morita M, Watanabe A, Morita M, Kakeji Y, Inoue F, Maehara Y. Colorectal cancer screening with odour material by

canine scent detection. Gut. 2011 Jun;60(6):814–9.

309. de Lacy Costello B, Amann A, Al-Kateb. A review of the volatiles from the healthy human body. J Breath Res. 2014;27:026001.

310. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten a, Tisch U, Haick H. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer. 2010 Aug 10;103(4):542–51.

311. Altomare DF, Di Lena M, Porcelli F, Trizio L, Travaglio E, Tutino M, Dragonieri S, Memeo V, de Gennaro G. Exhaled volatile organic compounds identify patients with colorectal cancer. Br J Surg. 2013 Jan;100(1):144–50.

312. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, Westenbrink E, Hodges P, Thomas MG, Chambers S, O'Connell N, Bailey C, Harmston C, Nwokolo CU, Bardhan KD, Covington J a. Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. PLoS One. 2014 Jan;9(9):e108750.

313. Buljubasic F, Buchbauer G. The scent of human diseases: a review on specific volatile organic compounds as diagnostic biomarkers. Flavour Fragr J. 2014 Oct 10;(August):n/a-n/a.

314. C. L. Silva, M. Passos JSC. nvestigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. Br J Cancer. 2011;105(12):1894–1904.

315. de Meij TG, Larbi I Ben, van der Schee MP, Lentferink YE, Paff T, Terhaar Sive Droste JS, Mulder CJ, van Bodegraven A a, de Boer NK. Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: proof of principle study. Int J Cancer. 2014 Mar 1;134(5):1132–8.

316. de Boer NKH, de Meij TGJ, Oort F a, Ben Larbi I, Mulder CJJ, van Bodegraven A a, van der Schee MP. The scent of colorectal cancer: detection by volatile organic compound analysis. Clin Gastroenterol Hepatol. 2014 Jul;12(7):1085–9.

317. Aggio RBM, Mayor A, Coyle S, Reade S, Khalid T, Ratcliffe NM, Probert CSJ. Freeze-drying: an alternative method for the analysis of volatile organic compounds in the headspace of urine samples using solid phase micro-extraction coupled to gas chromatography - mass spectrometry. Chem Cent J. 2016;10(1):9.

318. Khalid T, Aggio R, White P, De Lacy Costello B, Persad R, Al-Kateb H, Jones P, Probert CS, Ratcliffe N. Urinary Volatile Organic Compounds for the Detection of Prostate Cancer. PLoS One. 2015;10(11):e0143283.

319. Bond A, Vernon A, Reade S, Mayor A, Minetti C, Wastling J, Lamden K, Probert C. Investigation of Volatile Organic Compounds Emitted from Faeces for the Diagnosis of Giardiasis. J Gastrointest Liver Dis. 2015;24(3):281–6.

320. Aggio R, Villas-Bôas SG, Ruggiero K. Metab: an R package for high-throughput analysis of metabolomics data generated by GC-MS. Bioinformatics. 2011 Aug 15;27(16):2316–8.

321. Ihaka R GR. A language for data analysis and graphics. J Comput Graph Stat. 1996;5:299–314.

322. R Development Core Team (2009) R. A language and environment for statistical computing. R Foundation for Statistical Computing. 2009.

323. Xia, J., Mandal, R., Sinelnikov, I., Broadhurst, D., and Wishart DS. MetaboAnalyst 2.0 - a comprehensive server for metabolomic data analysis . Nucl Acids Res. 2009;40:W127–33.

324. Xia, J., Sinelnikov, I., Han, B., and Wishart D. MetaboAnalyst 3.0 - making metabolomics more meaningful . Nucleic acid Res. 2015;43:W251-7.

325. Ioannou GN, Perkins JD, Carithers RL. Liver Transplantation for Hepatocellular Carcinoma: Impact of the MELD Allocation System and Predictors of Survival. Gastroenterology. 2008;134(5):1342–51.

326. Llovet, JM, Briux, J. Early diagnosis and treatment of hepatocellular carcinoma. Baillieres Best Pract Res Clin Gastroenterol. 2000;14(6):991–1008.

327. Sinn DH, Cho JY, Gwak GY, Paik YH, Choi MS, Lee JH, Koh KC, Paik SW, Yoo BC. Different survival of barcelona clinic liver cancer stage C hepatocellular carcinoma patients by the extent of portal vein invasion and the type of extrahepatic spread. PLoS One. 2015;10(4):1–12.

328. Marrero JA, Fontana RJ BA. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. Hepatology. 2005;41:707–16.

329. Bruix J, Sherman M. Management of hepatocellular carcinoma: An update. Hepatology. 2011;53(3):1020–2.

330. Colombo M. Screening for cancer in viral hepatitis. Clin liver Dis. 2001;5:109–22.

331. Talseth BA, Meldrum C, Suchy J, Kurzawski G, Lubinski J, Scott RJ. Genetic polymorphisms in xenobiotic clearance genes and their influence on disease expression in hereditary nonpolyposis colorectal cancer patients. Cancer Epidemiol Biomarkers Prev. 2006;15(11):2307–10.

332. NTP. Report on Carcinogens, Thirteenth Edition. 2014;

333. The National Academy of Sciences. Review of the formaldehyde assessment in the National Toxicology Program 12th report on carcinogens. 2014.

334. Crowell PL, Elson CE, Bailey HH, Elegbede A, Haag JD GM. Human metabolism of the experimental cancer therapeutic agent d-limonene. Cancer Chemother Pharmacol. 1994;35(1):31–7.

335. Public Health England. Phenol. 2016.

336. Gamage N, Barnett A, Hempel N, Duggleby RG, Windmill KF, Martin JL, McManus ME. Human sulfotransferases and their role in chemical metabolism. Toxicol Sci. 2006;90(1):5–22.

337. Qin T, Liu H, Song Q, Song G, Wang H, Pan Y, Xiong F, Gu K, Sun G, Chen Z. The screening of volatile markers for hepatocellular carcinoma. Cancer Epidemiol Biomarkers Prev. 2010 Sep;19(9):2247–53.

338. Mochalski P, Sponring A, King J, Unterkofler K, Troppmair J, Amann A. Release and uptake of volatile organic compounds by human hepatocellular carcinoma cells (HepG2) in vitro. Cancer Cell Int. 2013 Jan;13(1):72.

339. Diren Beyoğlu1, Sandrine Imbeaud2, 3, Olivier Maurhofer1, Paulette Bioulac-Sage4, 5 J, Zucman-Rossi2, 3, Jean-François Dufour1 and JRI. TISSUE METABOLOMICS OF HEPATOCELLULAR CARCINOMA: TUMOR ENERGY METABOLISM AND THE ROLE OF

TRANSCRIPTOMIC CLASSIFICATION. Hepatology. 2014;58(1):229–38.

340. Office of National Statistics. 2011 UK Census. 2011.

341. Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics. 2006;2(4):171–96.

342. Bashir A, Miskeen AY, Bhat A, Fazili KM, Ganai B a. Fusobacterium nucleatum. Eur J Cancer Prev. 2015;24:373–385.

343. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. Nat Publ Gr. 2012;10(8):575–82.

344. zur Hausen H. The search for infectious causes of human cancers: where and why (Nobel lecture). Angew Chemie Int Ed. 2009;48:5798–5808.

345. de Martel C, Ferlay J, Franceschi S et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol. 2012;13:607–615.

346. Qin J, Li R, Raes J, Arumugam M, Burgdorf S, Manichanh C, Nielsen T, Pons N, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J, Hansen T, Paslier D Le, Linneberg A, Nielsen HB, Pelletier E, Renault P, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H. A human gut microbial gene catalog established by metagenomic sequencing. 2013;464(7285):59–65.

347. Garrett WS. Cancer and the microbiota. Science (80- ). 2015;348(6230).

348. Putze J, Hennequin C, Nougayrède JP, Zhang W, Homburg S, Karch H, Bringer MA, Fayolle C, Carniel E, Rabsch W, Oelschlaeger T a., Oswald E, Forestier C, Hacker J, Dobrindt U. Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. Infect Immun. 2009;77(11):4696–703.

349. Man SM. The clinical importance of emerging Campylobacter species. Nat Rev Gastroenterol Hepatol. 2011;8(12):669–85.

350. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G HY. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013;14:195–206.

351. Grivennikov SI KM. Inflammatory cytokines in cancer: tumour necrosis factor and interleukin 6 take the stage. Ann Rheumatol Dis. 2011;Suppl 1:i104-8.

352. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, El-Omar EM, Brenner D, Fuchs CS, Meyerson M, Garrett WS. Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. Cell Host Microbe. 2013;14(2):207–15.

353. Flanagan L, Schmid J, Ebert M, Soucek P, Kunicka T, Liska V, Bruha J, Neary P, Dezeeuw N, Tommasino M, Jenab M, Prehn JHM, Hughes DJ. Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. Eur J Clin Microbiol Infect Dis. 2014;33(8):1381–90.

354. Huycke, M. M., Abrams, V. & Moore DR. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. Carcinogenesis. 2002;23:529–536.

355. Wang, X. M. & Huycke MM. Extracellular superoxide production by Enterococcus faecalis promotes chromosomal instability in mammalian cells. Gastroenterology. 2007;132:551–561.

356. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. Escherichia coli induces DNA damage in vivo and triggers genomic instability in mammalian cells. Proc Natl Acad Sci U S A. 2010;107(25):11537–42.

357. Housseau, F. & Sears CL. Enterotoxigenic Bacteroides fragilis (ETBF)-mediated colitis in Min (Apc+/−) mice: a human commensal-based murine model of colon carcinogenesis. Cell cycle. 2010;9:3–5.

358. Warren RL, Freeman DJ, Pleasance S, Watson P, Moore R a, Cochrane K, Allen-Vercoe E, Holt R a. Co-occurrence of anaerobic bacteria in colorectal carcinomas. Microbiome. 2013;1:16.

359. Han YW. Oral Bacteria as Drivers for Colorectal Cancer. J Periodontol. 2014;(September):1155–7.

360. Belcheva A, Irrazabal T, Robertson SJ, Streutker C, Maughan H, Rubino S, Moriyama EH, Copeland JK, Kumar S, Green B, Geddes K, Pezo RC, Navarre WW, Milosevic M, Wilson BC, Girardin SE, Wolever TMS, Edelmann W, Guttman DS, Philpott DJ, Martin A. Gut microbial metabolism drives transformation of msh2-deficient colon epithelial cells. Cell. 2014;158(2):288–99.

361. Hellmich MR, Coletta C, Chao C, Szabo C. The Therapeutic Potential of Cystathionine β-Synthetase/Hydrogen Sulfide Inhibition in Cancer. Antioxid Redox Signal. 2014;22(5):424–48.

362. Baskar R and Bian J. Hydrogen sulfide gas has cell growth regulatory role. Eur J Pharmacol. 2011;656:5–9.

363. Corley D a, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni C a, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370(14):1298–306.

364. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370(14):2541.

365. Batty CA, Cauchi M, Lourenço C, Hunter JO, Turner C. Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer. PLoS One. 2015;10(6):e0130301.

366. Altomare DF, Di Lena M, Porcelli F, Travaglio E, Longobardi F, Tutino M, Depalma N, Tedesco G, Sardaro A, Memeo R, de Gennaro G. Effects of Curative Colorectal Cancer Surgery on Exhaled Volatile Organic Compounds and Potential Implications in Clinical Follow-up. Ann Surg. 2015;262(5):862–7.

367. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. Gut. 2015;64(8):1327–37.

368. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, Fedato C, Saugo M, Dei Tos AP. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut. 2015;64(5):784–90.

369. Halloran SP, Launoy G, Zappa M IA for R on C. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing.

Endoscopy. 2012;44:SE65-87.

370. Sung JJ, Ng SC, Chan FK, Chiu HM, Kim HS, Matsuda T, Ng SS, Lau JY, Zheng S, Adler S, Reddy N, Yeoh KG, Tsoi KK, Ching JY, Kuipers EJ, Rabeneck L, Young GP, Steele RJ, Lieberman D, Goh KL APWG. An updated Asia Pacific Consensus Recommendations on colorectal cancer screening. Gut. 2015;64(1):121–32.

371. Health D of. Improving outcomes: a strategy for cancer. 2011.

372. International Agency for Research on Cancer (IARC). International Agency for Research on Cancer (IARC). Agents Classified by the IARC Monographs. 2015.

373. Nougayrede JP, Homburg S, Taieb F, Boury M BE. Escherichia coli induces DNA double-strand breaks in eukaryotic cells. Science (80- ). 2006;313:848–851.

374. Bruce, A, Veral, l S, Hackett, CA & Wheatley R. Identification of volatile organic compounds (VOCs) from bacteria and yeast causing growth inhibition of sapstain fungi. Holzforschung. 58AD;193–198.

375. Weise T, Th??rmer A, Brady S, Kai M, Daniel R, Gottschalk G, Piechulla B. VOC emission of various Serratia species and isolates and genome analysis of Serratia plymuthica 4Rx13. FEMS Microbiol Lett. 2014;352(1):45–53.

376. Gao Z, Guo B, Gao R, Zhu Q, Qin H. Microbiota disbiosis is associated with colorectal cancer. Front Microbiol. 2015;6(February):1–9.

377. Wu N, Yang X, Zhang R, Li J, Xiao X, Hu Y, Chen Y, Yang F, Lu N, Wang Z, Luan C, Liu Y, Wang B, Xiang C, Wang Y, Zhao F, Gao GF, Wang S, Li L, Zhang H, Zhu B. Dysbiosis signature of fecal microbiota in colorectal cancer patients. Microb Ecol. 2013;66(2):462–70.

378. de Meij TG, Larbi IB, van der Schee MP. Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: proof of principle study. Int J cancer. 2014;134:1132–1138.

379. Probert CSJ, Jones PRH, Ratcliffe NM. A novel method for rapidly diagnosing the causes of diarrhoea. Gut. 2004 Jan;53(1):58–61.

380. Garner CE, Smith S, de Lacy Costello B, White P, Spencer R, Probert CSJ, Ratcliffe NM. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. FASEB J. 2007 Jun;21(8):1675–88.

381. Covington JA, Westenbrink EW, Ouaret N, Harbord R, Bailey C, O'Connell N, Cullis J, Williams N, Nwokolo CU, Bardhan KD, Arasaradnam RP. Application of a novel tool for diagnosing bile acid diarrhoea. Sensors. 2013;13(9):11899–912.

382. Soler AP, Miller RD, Laughlin K V., Carp NZ, Klurfeld DM, Mullin JM. Increased tight junctional permeability is associated with the development of colon cancer. Carcinogenesis. 1999;20(8):1425–31.

383. Ajibola O a, Smith D, Spaněl P, Ferns G a a. Effects of dietary nutrients on volatile breath metabolites. J Nutr Sci. 2013;2(2):e34.

384. Raman M, Ambalam P, Doble M. Probiotics and Bioactive Carbohydrates in Colon Cancer Management. 2016.

385. Shastri YM, Naumann M, Oremek GM, Hanisch E, Rösch W, Mössner J, Caspary WF, Stein JM.

Prospective multicenter evaluation of fecal tumor pyruvate kinase type M2 (M2-PK) as a screening biomarker for colorectal ecoplasia. Int J Cancer. 2006;119(11):2651–6.

386. Ahlquist DA. Fecal occult blood testing for colorectal cancer. Can we afford to do this? Gastroenterol Clin North Am. 1997;26(41–55).

387. Shah R, Jones E, Vidart V, Kuppen PJ, Conti JA FN. Biomarkers for early detection of colorectal cancer and polyps: systematic review. Cancer Epidemiol Biomarkers Prev. 2014;23:1712–28.

388. Shastri YM, Povse N, Schroder O SJ. Comparison of a novel fecal marker—fecal tumor pyruvate kinase type M2 (M2-PK) with fecal calprotectin in patients with inflammatory bowel disease: a prospective study. Clin Lab. 2008;54:389–90.

389. Chung-Faye G, Hayee B, Maestranzi S, Donaldson N, Forgacs I SR. Fecal M2-pyruvate kinase (M2-PK): a novel marker of intestinal inflammation. Inflamm Bowel Dis. 2007;13(11):1374–8.

390. Leen R, Seng-Lee C, Holleran G, O'Morain C, McNamara D. Comparison of faecal M2-PK and FIT in a population-based bowel cancer screening cohort. Eur J Gastroenterol Hepatol. 2014;26:514–8.

391. Parente F, Marino B, Ilardo A, Fracasso P, Zullo A, Hassan C, Moretti R, Cremaschini M, Ardizzoia A, Saracino I, Perna F VD. A combination of faecal tests for the detection of colon cancer: a new strategy for an appropriate selection of referrals to colonoscopy? A prospective multicentre Italian study. Eur J Gastroenterol Hepatol. 2012;24:1145–52.

392. . Ardizzoia, F. Parente, B. Marino, A. Ilardo, F. Perna, I. Saracino, A. Zullo, C. Hassan, P. Fracasso, R. A. Moretti, M. Cremaschini, D. Vaira GU. A combination of fecal tests for the detection of colon cancer: A new strategy for appropriate prioritization of referrals to colonscopy—A prospective Italian study. J Clin Oncol. 2011;29:(suppl; abstr 1580).

393. Noshirwani, KC, van Stolk RU, Rybicki LA BG. Adenoma size and number are predictive of adenoma recurrence: implications for surveillance colonoscopy. Gastrointest Endosc. 2000;51:433–7.

394. Atkin WS, Saunders BP. Surveillance guidelines after removal of colorectal adenomatous polyps. Gut. 2002;51 Suppl 5:V6–9.

395. Levin B, Lieberman D a., McFarland B, Andrews KS, Brooks D, Bond J, Dash C, Giardiello FM, Glick S, Johnson D, Johnson CD, Levin TR, Pickhardt PJ, Rex DK, Smith R a., Thorson A, Winawer SJ. Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline From the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology. 2008;134(5):1570–95.

396. Vishinkin R, Haick H. Nanoscale Sensor Technologies for Disease Detection via Volatolomics. Small. 2015;11(46):6142–64.

397. Moghadamyeghaneh Z, Hanna MH, Alizadeh RF, Carmichael JC, Mills S, Pigazzi A, Stamos MJ. Contemporary management of anastomotic leak after colon surgery: assessing the need for reoperation. Am J Surg. 2015;211(6):1–9.

398. Sammour T, Hayes IP, Jones IT, Steel MC, Faragher I, Gibbs P. Impact of anastomotic leak on recurrence and survival after colorectal cancer surgery: a BioGrid Australia analysis. ANZ J Surg. 2016 Jun 3. doi: 10.1111/ans.13648;

# Appendix 1

# Ethical approval

**NRES Committee South West - Central Bristol**
Whitefriars
Level 3, Block B
Lewin's Mead
Bristol BS1 2NT
Email: nrescommittee.southwest-bristol@nhs.net

Telephone: 0117 342 1335
Fax:0117 342 0445

08 December 2014

Prof Chris Probert
Professor of Gastroenterology
University of Liverpool
Dept of Gastroenterology
University of Liverpool
Liverpool
L69 3GE

Dear Prof Probert

**Study title:** **The role of faecal and urinary Volatile Organic Compounds (VOCs) as a non-invasive diagnostic biomarker of colorectal cancer.**
**REC reference:** **14/SW/1162**
**IRAS project ID:** **164287**

Thank you for your email of 8 December 2014. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 21 November 2014

**Documents received**

The documents received were as follows:

| *Document* | *Version* | *Date* |
|---|---|---|
| Participant consent form [Consent form] | 2 | 25 November 2014 |
| Participant information sheet (PIS) | 2 | 25 November 2014 |

**Approved documents**

The final list of approved documentation for the study is therefore as follows:

| *Document* | *Version* | *Date* |
|---|---|---|
| Letter from sponsor [Intention to sponsor] | 1 | 10 November 2014 |
| Non-validated questionnaire [Questionnaire] | 1 | 10 November 2014 |
| Participant consent form [Consent form] | 2 | 25 November 2014 |
| Participant information sheet (PIS) | 2 | 25 November 2014 |

| | | |
|---|---|---|
| REC Application Form [REC_Form_11112014] | | 11 November 2014 |
| Referee's report or other scientific critique report [Peer review] | 1 | 10 November 2014 |
| Research protocol or project proposal [Protocol] | 1 | 10 November 2014 |
| Summary CV for Chief Investigator (CI) [CI CV] | 1 | 10 November 2014 |

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

| | |
|---|---|
| **14/SW/1162** | **Please quote this number on all correspondence** |

Yours sincerely

**Naazneen Nathoo**
**REC Manager**

E-mail: nrescommittee.southwest-bristol@nhs.net

*Copy to:* *Mr Alex Astor*
*Sue Charnock, Royal Liverpool University Hospital*

NHS
Health Research Authority

**NRES Committee South East Coast - Brighton & Sussex**
Health Research Authority
Ground Floor, Skipton House
80 London Road
London
SE1 6LH

Telephone:
Fax:

11 May 2015

Prof Chris Probert
Professor of Gastroenterology
University of Liverpool
Dept of Gastroenterology
University of Liverpool
Liverpool
L69 3GE

Dear Prof Probert

| | |
|---|---|
| **Study title:** | **The role of urinary volatile organic compounds as a non-invasive diagnostic biomarker for hepatocellular carcinoma** |
| **REC reference:** | **15/LO/0836** |
| **Protocol number:** | **UoL001123** |
| **IRAS project ID:** | **173809** |

Thank you for your letter responding to the Proportionate Review Sub-Committee's request for changes to the documentation for the above study.

The revised documentation has been reviewed and approved by the sub-committee.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager Mrs. Alison O'Kane , NRESCommittee.SECoast-BrightonandSussex@nhs.net. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

**Confirmation of ethical opinion**

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

**Conditions of the favourable opinion**

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

*Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.*

*Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.*

*Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.*

*For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.*

*Sponsors are not required to notify the Committee of approvals from host organisations.*

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from NRES. Guidance on where to register is provided on the HRA website.

**It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).**

**Ethical review of research sites**

The favourable opinion applies to all NHS sites taking part in the study, subject to management

permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" above).

**Approved documents**

The documents reviewed and approved by the Committee are:

| *Document* | *Version* | *Date* |
|---|---|---|
| IRAS Checklist XML [Checklist_07052015] | | 07 May 2015 |
| Letter from sponsor [Sponsor letter] | 1 | 07 April 2015 |
| Non-validated questionnaire [Questionnaire] | 1 | 02 February 2015 |
| Participant consent form [consent form] | 1 | 02 February 2015 |
| Participant information sheet (PIS) [PIS] | 2 | 05 May 2015 |
| REC Application Form [REC_Form_07052015] | | 07 May 2015 |
| Referee's report or other scientific critique report [Peer review] | 1 | 25 February 2015 |
| Research protocol or project proposal [Protocol] | 2 | 05 May 2015 |
| Summary CV for Chief Investigator (CI) [CV for CI] | 1 | 01 February 2015 |

**Statement of compliance**

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

**After ethical review**

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance

We are pleased to welcome researchers and R & D staff at our NRES committee members'

training days – see details at http://www.hra.nhs.uk/hra-training/

| 15/LO/0836 | **Please quote this number on all correspondence** |
|---|---|

With the Committee's best wishes for the success of this project.

Yours sincerely

Alison O'Kane

**Chair**

Email: NRESCommittee.SECoast-BrightonandSussex@nhs.net

Enclosures: *"After ethical review – guidance for researchers"*

*Copy to:* *Mr Alex Astor*

*Mrs Sue Charnock, Royal Liverpool and Broadgreen University Hospital Trust*

# Appendix 2

# Patient information sheets and consent forms

# A non-invasive diagnostic indicator for Colorectal Cancer

**An invitation to you**

You are being invited to take part in a research study. **You DO NOT have to take part** and your treatment WILL NOT BE AFFECTED in any way by any involvement, or otherwise, in this project. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please read the following information carefully and discuss it with friends, relatives or your GP/ward doctor as you wish. Please do not hesitate to contact us if there is any aspect of the research that is not clear or if you would like more information. The purpose of this document is to invite you to take part in medical research.

**Purpose of this study**

We intend to collect the gas that makes the smell from your stool and urine samples. Your samples will be used to investigate whether there are any changes in the chemicals contained in the faeces and urine which may be used to diagnose colorectal (large bowel) cancer or polyps (a potentially early form of the cancer). We are aware that patients with a wide variety of bowel and liver disorders experience unpleasant faecal odour. Recent evidence also suggests that dogs can detect a change in the smell from their owners, if they develop cancer.

We have already found several chemicals that may be used to determine when the stool contains various bacteria and viruses. We will analyse the faecal and urine odours to determine whether particular smells are linked to colorectal cancer and colonic polyps.

### *Nature of the research*

The chemicals making up faecal and urinary odours will be analysed from the space in the closed tube above the samples. The gas, stool, urine contents and any bacteria grown from the samples will not be retained after the completion of the research project. All samples will be entirely destroyed by incineration upon completion of the project.

**Why have you been chosen to participate?**

There a number of different types of patients being asked to participate in this study.

You may have been selected as you have a positive Faecal Occult Blood test. This means that the presence of a small amount of blood has been detected in your stool. Blood in the stool can be one sign of colorectal (large bowel) cancer or colonic polyps and you have been advised to attend for a colonoscopy (camera test) to explore this possibility further.

Or

You are having a colonoscopy for known polyp screening/surveillance

Or

You are having a colonoscopy for a change in bowel habit or iron deficiency anaemia. Please note that if your colonoscopy is normal you will be included in the study as our control (normal) group.

Or

You are due to undergo an operation for a known/possible colorectal cancer. Or you are under the care of the colorectal team.

We would like a sample of your stool and urine.

**What is required of me?**

Please collect a sample of stool and urine before you take your bowel preparation (laxative) medication, using the equipment provided, and place it in the dedicated pots in the pack. Once you have given us the samples, we will ask you to fill in a form to give some more information about your age, ethnic origin and whether the stool sample comes from a stool that is normal for you. Please label the pots with your name and date of birth, enclose the sample pot in the small plastic bag provided and hand both the pots and the completed form in to the screening nurses at your colonoscopy.

**Who is organising the study?**

Professor Chris Probert, Professor in Gastroenterology, leads this study. Dr Ashley Bond, Clinical Research Fellow, will handle sample collection, data handling and experiments on the samples to analyse the odour. Both are doctors employed by the University of Liverpool and Royal Liverpool University Hospital

**What will happen to me if I take part?**

After you pass the two specimens into the sample tubes and provide this to us at your colonoscopy, this will be frozen and we will analyse the odour to identify its chemical composition. No other input is required from you. Nothing additional will happen to you as part of this study.

**Do I have to donate this faecal sample?**

No. You are completely free to decline donation - this will not affect your treatment.

**Are there any risks or disadvantages in taking part in this study?**

No.

**What are the possible benefits of taking part?**

This research on faecal and urinary odours will help us to identify chemicals that will hopefully allow earlier diagnosis of colorectal cancer and therefore increase the potential for cure. There will be no direct benefit to you.

**Is my doctor being paid for including me in this study?**

No.

**What if something goes wrong?**

There are no potential side effects of using your faecal and urine samples in this study.

**Confidentiality - who will know I am taking part in the research?**

Prof. Probert, Dr Bond and the management of the NHS Bowel Cancer Screening Programme. All of the individuals mentioned above are bound by NHS confidentiality requirements and the Data Protection Act.

**Confidentiality – sample storage and data handling**

Details of your age, sex and selected questions which will have relevance to the interpretation of your sample results will be gathered on the attached form. These data, along with the sample, will then be catalogued and stored confidentially and anonymously. All data will be disposed of at the end of the project. Please be advised we will follow up your medical records for comparison with the VOCS results

**Confidentiality - anonymity**

Your sample will be catalogued with unique identifier. This can only be traced back to you by Professor Probert or Dr. Bond, who will have access to the master list.

**Retention of tissue and faecal and samples**

Your sample will be used only for the research specified and will be destroyed by incineration upon completion of the research project. No tissue is being retained or stored for any other purpose.

**Local Ethical Committee Approval**

All research in the NHS is looked at by independent group of people called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the Royal Liverpool University Hospital and University of Liverpool Research and Development Departments, the local (Mersey) Research Ethics Committee and the Colorectal Cancer National Screening Programme to ensure that it meets the NHS national standards for research involving human patients. South

West-Central Bristol REC have reviewed and approved the study.

**What will happen to the results of the study?**

We will use the data obtained from your samples to try to devise a test for the early diagnosis of colorectal (large bowel) cancer. The result of this study will be published in an appropriate journal, all results will be anonymised.

**Further information**

This may be obtained from:

Department of Gastroenterology,

University of Liverpool

Tel: 0151 794 6822

abond@liverpool.ac.uk

PALS – Patient Advice and Liaison Service (independent of us)

Tel: 0800 218 2333

Thank you for reading this.

We hope that you will be able to help us with our research.

CS Probert and AD Bond

General history about your health *(please write your answers next to the questions)*:

1. Are you a current smoker?

2. Do you have any past or current bowel problems?

3. Are you a vegetarian?

4. Have you taken any antibiotics at any point over the last 6 months?

5. Approximately how many units of alcohol do you drink each week? (1 pint of lager or 1 normal 175ml glass of red wine is approx 2.5 units)

6. Place of birth/ethnicity

7. Regular medication

*CONSENT FORM The role of faecal and urinary Volatile Organic Compounds (VOCs) as a non-invasive diagnostic biomarker of colorectal cancer.*

1. I confirm that I have read and understand the information sheet for the above study (Version 2, date 25/11/14) and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my medical care or legal rights being affected. ☐

3. I understand that relevant sections of my medical notes, and data collected during the study, may be looked at by individuals from regulatory authorities or from the NHS Trust where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

4. I agree to take part in the above study. ☐

# A non-invasive diagnostic indicator for Hepatocellular carcinoma

**An invitation to you**

You are being invited to take part in a research study. **You DO NOT have to take part** and your treatment WILL NOT BE AFFECTED in any way by any involvement, or otherwise, in this project. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please read the following information carefully.. . The purpose of this document is to invite you to take part in medical research.

**Purpose of this study**

We intend to collect the gas that makes the smell from your urine samples. Your samples will be used to investigate whether there are any changes in the chemicals contained in the urine which may be used to diagnose Hepatocellular (liver) cancer...

We have already found several chemicals that may be used to determine when the urine contains a scent associated with liver disease and liver cancer. We will analyse the urine odour to determine whether particular smells are linked to liver cancer.

## *Nature of the research*

The chemicals making up urine odour will be analysed from the space in the closed tube above the urine sample. The gas and urine contents from the samples will not be retained after the completion of the research project. All samples will be entirely destroyed by incineration upon completion of the project.

**Why have you been chosen to participate?**

You have been selected to participate as you have one of the one of the following conditions:

1) Chronic liver disease without cirrhosis
2) Chronic liver disease with cirrhosis
3) A diagnosis of Hepatocellular carcinoma

It is anticipated that we will use those without a cancer diagnosis as our control group

**What is required of me?**

All we require from you is a urine sample today during your clinic appointment and completion of the brief questionnaire. That will end your required input and will in no way affect your ongoing management.

**Who is organising the study?**

Professor Chris Probert, Professor in Gastroenterology, leads this study. Dr Ashley Bond, Clinical Research Fellow, will handle sample collection, data handling and experiments on the sample to analyse the odour. Both are doctors employed by the University of Liverpool and Royal Liverpool University Hospital. There will also be assistance from Dr Tim Cross at the Royal Liverpool University Hospital and Dr Nick Stern at University Hospital Aintree. Dedicated nurse specialist and research nurse will discuss sample collection with you at each site.

**What will happen to me if I take part?**

After you provide the urine sample during the clinic appointment, this will be frozen and we will analyse the odour to identify its chemical composition. No other input is required from you. Nothing additional will happen to you as part of this study.

**Do I have to donate this urine sample?**

No. You are completely free to decline donation - this will not affect your treatment.

**Are there any risks or disadvantages in taking part in this study?**

No.

**What are the possible benefits of taking part?**

This research on urine odours will help us to identify chemicals that will hopefully allow earlier diagnosis of liver cancer and therefore increase the potential for cure. There will be no direct benefit to you.

**Is my doctor being paid for including me in this study?**

No.

**What if something goes wrong?**

There are no potential side effects of using your urine sample in this study.

**Confidentiality - who will know I am taking part in the research?**

Prof. Probert, Dr Bond. All of the individuals mentioned above are bound by NHS confidentiality requirements and the Data Protection Act.

### *Confidentiality – sample storage and data handling*

Samples will be allocated an anonymous study number and all data will be anonymous. Only clinical information relevant to the study will be looked at by the investigators.All data will be disposed of at the end of the project.

### *Confidentiality - anonymity*

Your sample will be catalogued with unique identifier. This can only be traced back to you by Professor Probert or Dr. Bond, who will have access to the master list.

### *Retention of urine samples*

Your sample will be used only for the research specified and will be destroyed by incineration upon completion of the research project. No tissue is being retained or stored for any other purpose.

**Local Ethical Committee Approval**

All research in the NHS is looked at by independent group of people called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by the NRES Committee South East Coast Brighton and Sussex.

**What will happen to the results of the study?**

We will use the data obtained from your samples to try to devise a test for the early diagnosis of hepatocellular carcinoma (liver cancer). We intend to seek publication of the results in academic journals, you would have full access to these potential results.

**Further information**

This may be obtained from:

Dr Ashley Bond

Department of Gastroenterology,

University of Liverpool

Tel: 0151 794 6822

Email: abond@liverpool.ac.uk

PALS – Patient Advice and Liaison Service (independent of us)

Tel: 0800 218 2333

Thank you for reading this.

We hope that you will be able to help us with our research.

CS Probert and A Bond

General history about your health *(please write your answers next to the questions)*:

1. Are you a current smoker?

2. Are you a vegetarian?

3. Approximately how many units of alcohol do you drink each week? (1 pint of lager or 1 normal 175ml glass of red wine is approx 2.5 units)

4. Place of birth/ethnicity

5. Regular medication

CONSENT FORM

*The role of urinary volatile organic compounds as a non-invasive diagnostic biomarker for hepatocellular carcinoma.* Principal Investigator: Professor Chris Probert Please initial box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my medical care or legal rights being affected. ☐

3. I understand that relevant sections of my medical notes, and data collected during the study, may be looked at by individuals from regulatory authorities or from the NHS Trust where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

4. I agree to take part in the above study. ☐

_____ _____ _____

Patients name Date Signature

_____ _____ _____

Name of person taking consent Date Signature

# Appendix 3

# Supporting publications

ORIGINAL PAPER

Available from: URL: .....
DOI: http://dx.doi.org/.....

# Correlation between Faecal Tumour M2 Pyruvate Kinase and Colonoscopy for the Detection of Adenomatous Neoplasia in a Secondary Care Cohort

**Ashley D. Bond[1], Michael D. Burkitt[1], David Sawbridge[1], Bernard M. Corfe[2,3], Chris S. Probert[1]**

1) Unit of Gastroenterology Research, Dept of Molecular and Cellular Physiology, Institute of Translational Medicine, University of Liverpool, Liverpool;
2) Molecular Gastroenterology Research Group, Department of Oncology;
3) Insigneo Institute for *in silico* Medicine, University of Sheffield Sheffield, UK

***Address for correspondence:***
Dr **Ashley Bond**
Dept of Translational Medicine,
University of Liverpool,
Liverpool, UK
abond@liverpool.ac.uk

Received: 15.12.2016
Accepted: 01.02.2016

## ABSTRACT

**Background & Aims**: Colorectal cancer screening programmes that target detection and excision of adenomatous colonic polyps have been shown to reduce colorectal cancer related mortality. Many screening programmes include an initial faecal occult blood test (FOBt) prior to colonoscopy. To refine the selection of patients for colonoscopy other faecal-based diagnostic tools have been proposed, including tumour M2-pyruvate kinase (tM2-PK). To determine whether tM2-PK quantification may have a role in diverse settings we have assessed the assay in a cohort of patients derived from both the England bowel cancer screening programme (BCSP) and symptomatic individuals presenting to secondary care.

**Method**. Patients undergoing colonoscopy provided faecal samples prior to bowel preparation. Faecal tM2-PK concentrations were measured by ELISA. Sensitivity, specificity, positive predictive value, negative predictive value and ROC analyses were calculated.

**Results.** Ninety-six patients returned faecal samples: 50 of these with adenomas and 7 with cancer. Median age was 68. Median faecal tM2-PK concentration was 3.8 U/mL for individuals without neoplastic findings at colonoscopy, 7.7 U/mL in those with adenomas and 24.4 U/mL in subjects with colorectal cancer (both, p=0.01). ROC analysis demonstrated an AUROC of 0.66 (sensitivity 72.4%, specificity 48.7%, positive predictive value 67.7%, negative predictive value 36.7%). Amongst BCSP patients with a prior positive FOBt faecal tM2-PK was more abundant (median 6.4 U/mL, p=0.03) and its diagnostic accuracy was greater (AUROC 0.82).

**Conclusion**. Our findings confirm that faecal tM2-PK ELISA may have utility as an adjunct to FOBt in a screening context, but do not support its use in symptomatic patients.

**Key words:** colorectal cancer – faecal tM2-PK – surveillance – screening – adenomatous polyps.

**Abbreviations**: BCSP: Bowel cancer screening programme; EMR: Endoscopic mucosal resection; FAP: Familial adenomatous polyposis; FOBt: Faecal occult blood testing; NHS: National Health Service; tM2-PK: tumour M2-pyruvate kinase.

## INTRODUCTION

Colorectal cancer is a leading cause of mortality and morbidity worldwide, with an estimated European incidence of 43.5 per 100,000 in 2012 and mortality of 19.5 per 100,000 [1]. UK lifetime risk is 1 in 15 for men or 1 in 19 for women [2]. Across Europe, colorectal cancer is the second most common cause of cancer related mortality [1]. Colorectal cancer carries a significant financial burden for the National Health Service (NHS), with a mean annual cost of £ 12,000 and £8,800 for those with rectal and non-rectal colon cancer, respectively [3].

A number of well-characterised genetic conditions increase the risk of colorectal cancer, such as Familial Adenomatous Polyposis (FAP); however, >90 % of colon cancers arise sporadically from adenomatous polyps. Adenomatous polyps develop over a 10- to 15-year period through the adenoma-carcinoma pathway [4]. The potential for malignant transformation combined with the long lag-time, during which identifiable, treatable, premalignant lesions exist, makes colorectal cancer an excellent candidate for population-based screening. This is supported by evidence that early diagnosis improves outcomes: patients presenting with early, Duke's A stage, disease have a 93.2 % chance of 5-year survival, whilst individuals presenting with advanced, Duke's C or D, disease have only 47.7 % and 6.6 % 5-year survivals, respectively [5].

Europe-wide guidance for bowel cancer screening recommends a faecal occult blood test (FOBt) between the ages of 50-74, followed by colonoscopy if FOBt is positive [6]. Screening programme vary across the UK; for example in Scotland, it is performed from 50 to 74 years of age with optional opt-in after 75 years. Currently in England all people between 60-74 years of age are sent for a FOBt; there is an optional opt-in after 75. Positive FOBt is usually followed by colonoscopy. Two forms of FOBt (immunological, i-FOBt and guaiac, g-FOBt) are currently used for screening purposes worldwide [7]. Currently in England, g-FOBt is used in the initial part of the screening process.

The g-FOBt uses guaiac (derived from wood resin of Guajacum trees) to detect haem in faeces. This test is particularly prone to false positive results because of ingestion of foods including red meat, carrots, potatoes and figs [8]. The g-FOBt identifies a daily blood loss of approximately 10 mL. Immunological FOBt can provide both qualitative and quantitative results. Qualitative i-FOBt is an instant and objective test which utilizes immuno-chromatography to detect occult blood loss in feces. However, most i-FOBts are quantitative measures of fecal haemoglobin using automated immunoturbidimetry. i-FOBt is specific to human haemoglobin and only reacts with the intact haemoglobin molecule. Therefore, i-FOBt requires no dietary restriction. In 2010, a meta-analysis reported a sensitivity (67%) and specificity (8 %) for immunological FOBt compared with 54% and 80% for g-FOBt, for the detection of colorectal cancer and pre-cancerous neoplasia [7]. The low sensitivity of g-FOBt has led to criticism of its use for population-based screening [9].

An alternative faecal-based screening tool is tumour M2-pyruvate kinase (tM2-PK) [9, 10]. tM2-PK is an isoenzyme of pyruvate kinase; it is involved in glycolysis and catalyzes the ATP-producing conversion of phosphoenolpyruvate to pyruvate. During tumourigenesis other isoenzymes of pyruvate kinase are lost and there is an increased expression of the M2 subtype [11]. In neoplastic tissue, M2-PK is mainly found in the dimeric form, rather than tetrameric, driven by modifications by a number of oncoproteins. Dimeric M2-PK is more abundant in cancerous states, including ovarian, lung and colorectal [12], and has therefore been described as tumour M2-PK (tM2-PK). Abundance of tM2-PK can be quantified by a sandwich enzyme-linked immunosorbent assay (ELISA, ScheBo Biotech AG, Giessen, Germany). A number of studies have demonstrated increased abundance of plasma tM2-PK and correlation with stage of melanoma, thyroid, breast, lung, kidney, oesophageal, gastric, pancreatic, colorectal, ovarian, cervical and renal cell cancer [9, 13, 14]. Increased plasma concentration of tM2-PK has also been described in other colonic pathology, including diverticulosis and inflammatory bowel disease. However, a meta-analysis in 2012 concluded that the use of faecal tM2-PK in colorectal cancer screening would "close a clinical gap", because of its high sensitivity and specificity [12].

In the current study, we have assessed the faecal tM2-PK ELISA in a novel secondary care cohort of patients derived from (1) the England Bowel Cancer Screening Programme (BCSP), being conducted locally, and (2) symptomatic patients attending the Royal Liverpool University Hospital in an attempt to characterise whether this assay may have a role in diverse settings. Such utility may include its use as a pre-screen to colonoscopy in order to stratify risk and allow for service planning or as an adjunct to the current BCSP.

## MATERIAL AND METHODS

### Patient recruitment

Patients undergoing colonoscopy for the investigation of iron deficiency anaemia, polyp surveillance, intervention for known polyps e.g. endoscopic mucosal resection (EMR), change in bowel habit and investigation of abnormal radiology results were eligible for inclusion. These patients were not part of a dedicated screening programme. Those participating in the England BCSP being conducted locally i.e. after a positive g-FOBt, were also eligible for inclusion. Recruitment was directed at these patients, rather than all patients attending for colonoscopy in order to maximise the burden of adenoma/carcinoma pathology. Collection kits were sent to patients prior to their intended colonoscopy. They were asked to produce their sample within 48 hours of their colonoscopy, but before starting bowel preparation medication. Patients followed the usual dietary modifications required prior to colonoscopy. Samples were returned at the time of colonoscopy: thus, all samples were frozen within 48 hours of egestion. Colonoscopy results, including any histological findings, were recorded. Patients were then categorised as having malignant, adenomatous or non-neoplastic ("normal") results. In those with adenomatous disease location, size and number of polyps were recorded. Polyps were assigned to the adenoma group only when this was confirmed by histology. Hyperplastic polyps were classified as "normal".

### Ethical consideration

Full ethical committee approval was given by the NRES Committee South West - Central Bristol, reference 14/SW1162.

### Sample handling and statistical analysis

Faecal samples were delivered by patients when they attended for colonoscopy and were immediately transferred to the laboratory and frozen at -20°C in the same glass headspace vials (Supelco, USA) in which they were collected, until analysis. Storage was in line with the ELISA manufacturer's instructions. tM2-PK concentration was measured with a commercially available sandwich ELISA (ScheBo Biotech AG, Giessen, Germany). Assays were performed in our university-based research laboratory by a researcher who had undergone specific training and validation provided by the assay manufacture. This analytical performance of each plate was assessed by interpolated control calculation using a sample of known tM2-PK abundance. If a deviation >15 % from expected concentration was identified assay failed QC and was repeated. All analyses were performed by an investigator blinded to the patients' diagnoses, under standardised conditions. Graphpad Prism 6 was used to compare pathological groups by the two tailed *t* test and Kruskal-Wallis and Dunn's post-hoc analyses (p value significant <0.05), and to determine the area under ROC

curves, sensitivity, specificity, positive predictive value and negative predictive values. Further testing was performed using a binomial method McNemar's test.

## RESULTS

### Patients' characteristics

Ninety-six patients returned samples within the allotted time. Thirty-six patients were participating in the BCSP and, therefore, had a prior positive g-FOBt. The remaining 60 were symptomatic patients outside of the BCSP and had an unknown FOBt status: the indications for their procedures were iron deficiency anaemia (n=18), change in bowel habit (n=15), surveillance or intervention for known polyps (n=20), family history of colorectal cancer and abnormal radiology (n=7). Fifty-two of the cohort were male and 44 female. The median age was 68 years. Faecal tM2-PK levels were significantly higher in males [15.37 U/mL (95% CI 10.8-19.8) vs 8.034 U/mL (95% CI 5.01-10.9), p=0.01 by 2-tailed Student's *t*-test] (Fig. 1A). To characterise whether patient's age influenced faecal tM2-PK concentration the cohort was divided into quintiles according to their age at the time of sampling. No significant differences in mean faecal tM2-PK were identified between any two quintiles (Kruskal-Wallis 1-way ANOVA and Dunn's post-hoc analysis; Fig. 1B).

### tM2-PK concentration is elevated in patients with adenomatous and malignant disease

Median faecal tM2-PK concentration in individuals with non-neoplastic colonoscopy results was 3.8 U/mL, in those with either adenomatous or malignant disease the concentration was significantly greater (median 7.7 U/mL and 24.4 U/mL, respectively: p=0.01 by Kruskal Wallis 1-way ANOVA and Dunn's post-hoc analysis, Fig. 1C).

### Diagnostic utility

To assess the suitability of the faecal tM2-PK ELISA as a screening test in our cohort, we categorised patients into those with non-neoplastic disease and those with a burden of neoplastic disease by colonoscopy. We performed a ROC analysis on the basis of these two groups and showed that the assay may have limited overall utility with an area under the ROC curve of 0.66 (p=0.006) (Fig. 2A).

The ELISA manufacturer proposed a cut-off value of 4.0 U/mL. Adopting this as a cut-off to define a positive tM2-PK test, the assay's sensitivity to detect adenoma or carcinoma was 72.4% (95%CI; 59.1-83.3%), specificity 48.7% (95%CI; 32.4-65.2%), positive predictive value 67.7% (95%CI; 54.7-79.2%) and negative predictive value 36.7 % (95%CI; 36.7-71.2%). Application of McNemar's test did not show a significant difference between the tM2-PK results and the presence of neoplasia at colonoscopy (p=0.73), suggesting a limited relationship between the faecal tM2-PK value and findings at colonoscopy (Appendix 1). There was a false positive rate of 20.6% across the 96 samples. In total 37% of samples were incorrectly categorised by tM2-PK testing.

The same statistical analysis was performed using a series of different threshold levels (Table I and Fig. 2B). The optimal cut-off value identified by the highest McNemar's test was at 4.8: this gave a sensitivity of 69% and a specificity of 56%.

Of the 50 patients with adenomatous disease, 22 were found to have a single polyp greater than 1cm in size. The median faecal tM2-PK reading for this group was 5.7 U/mL (interquartile range 1.3-19.5). Of these patients, 59% (n=13) had a faecal tM2-PK result above the 4.0 U/mL cut-off, representing a true positive result by the assay manufacturer's criteria. There was no significant difference in the tM2-PK concentration detectable in the faeces of patients with polyps >1cm or <1cm, p=0.63 (Fig. 3A). There was also no significant difference in the tM2-PK concentration in faeces of those with <5 or ≥5 adenomatous polyps (Fig. 3B).

### Assessment of role as adjunct to FOBt in BCSP

When we isolated those individuals (n=36) referred for colonoscopy as part of the BCSP, those with neoplasia (n=25) were found to have significantly greater faecal tM2-PK concentration than those without neoplastic pathology (Fig. 4A, p=0.03). A macroscopically normal colonoscopy was seen in all BCSP patients who did not otherwise have neoplastic disease (n=11). ROC analysis of the BCSP patients alone demonstrated an AUROC of 0.82 (p=0.002) (Fig. 4B). Sixty one percent of patients with a prior positive fFBOt also had a positive tM2-PK. Of the BCSP with no neoplasia, 72% had a negative tM2-PK, whilst 76% of those with neoplasia had a positive tM2-PK.

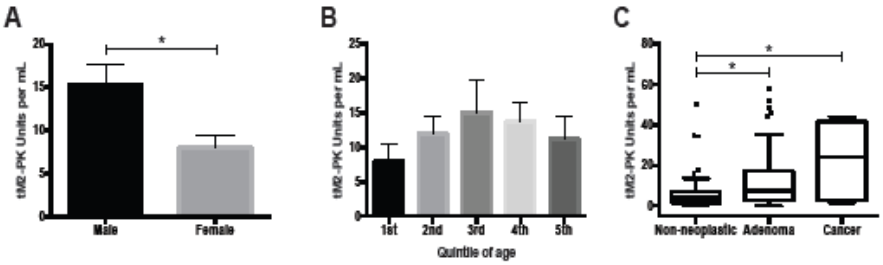



**Fig. 1.** A) Bar chart representing concentration of tM2-PK according to the gender of patient (*p<0.05, by Student's *t*-test). B) Bar chart representing concentration of tM2-PK according to the age of patients separated by quintiles. C) Box and whisker plot demonstrating tM2-PK concentration stratified by diagnosis at colonoscopy. (*p<0.05, by Kruskal-Wallis 1-way ANOVA and Dunn's post-hoc analysis)

**Table I.** Table showing the sensitivity and specificity of different faecal tM2-PK thresholds for the study cohort, with binomial method McNemar values for each of these thresholds.

| Threshold | Sensitivity % | 95% CI | Specificity % | 95% CI | Likelihood ratio | McNemar Binomial p-value |
|---|---|---|---|---|---|---|
| 3 | 75.86 | 62.83 to 86.13 | 41.03 | 25.57 to 57.90 | 1.286 | 0.2559 |
| 3.2 | 74.14 | 60.96 to 84.75 | 41.03 | 25.57 to 57.90 | 1.257 | 0.2559 |
| 3.4 | 74.14 | 60.96 to 84.75 | 43.59 | 27.81 to 60.38 | 1.314 | 0.324 |
| 3.6 | 74.14 | 60.96 to 84.75 | 43.59 | 27.81 to 60.38 | 1.314 | 0.324 |
| 3.8 | 74.14 | 60.96 to 84.75 | 48.72 | 32.42 to 65.22 | 1.446 | 0.6177 |
| 4 | 72.41 | 59.10 to 83.34 | 48.72 | 32.42 to 65.22 | 1.412 | 0.7359 |
| 4.2 | 72.41 | 59.10 to 83.34 | 48.72 | 32.42 to 65.22 | 1.412 | 0.7359 |
| 4.4 | 72.41 | 59.10 to 83.34 | 51.28 | 34.78 to 67.58 | 1.486 | 0.7359 |
| 4.6 | 72.41 | 59.10 to 83.34 | 53.85 | 37.18 to 69.91 | 1.569 | 0.8642 |
| 4.8 | 68.97 | 55.46 to 80.46 | 56.41 | 39.62 to 72.19 | 1.582 | 0.8679 |
| 5.0 | 67.24 | 53.66 to 78.99 | 56.41 | 39.62 to 72.19 | 1.543 | 0.7428 |
| 5.2 | 65.52 | 51.88 to 77.51 | 56.41 | 39.62 to 72.19 | 1.503 | 0.7428 |

This difference was not observed in the non-BCSP group; 29 of whom were classified as non-neoplastic; of these 7 had IBD and/or diverticulosis. Within this subset, 6 had elevated (false positive) tM2-PK concentration. The remaining 22 subjects had a macroscopically normal colonoscopy, of whom 11 received a true negative faecal tM2-PK result, whilst 11 had falsely positive results. Within this subset AUROC was not significantly different to 0.5 (AUROC 0.51, p=0.87)(Fig. 4C).

## DISCUSSION

The faecal tM2-PK concentration observed in this study demonstrated an increasing trend as pathology progressed along the adenoma-carcinoma pathway, supporting potential utility for the assay to distinguish between absence and presence of neoplastic disease.

Our cohort contained more males, with a median age of 68, therefore representative of a potential colorectal cancer

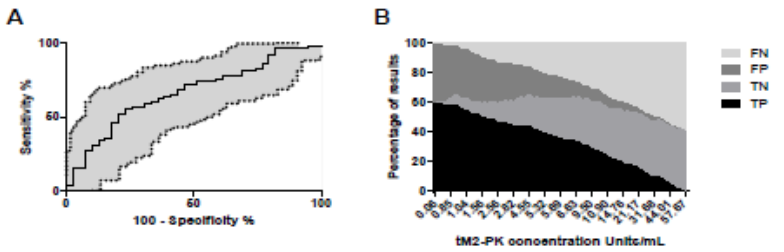



**Fig. 2.** A) ROC curve showing utility of faecal tM2-PK ELISA for detection of adenomatous disease in this cohort, shaded area represents 95%CIs. B) Surface plot demonstrating the variation in diagnostic accuracy as the diagnostic threshold concentration changes FN: false negative, FP: false positive, TN: true negative, TP: true positive.

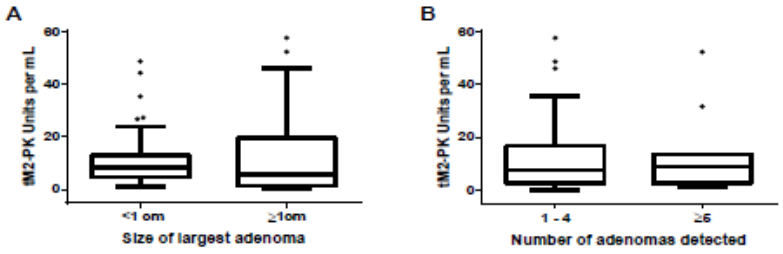



**Fig. 3.** A) Box and whisker plot showing faecal tM2-PK concentration stratified by size of largest adenoma. B) Box and whisker plot of faecal tM2-PK concentration stratified by the number of adenomas indentified in each patient.

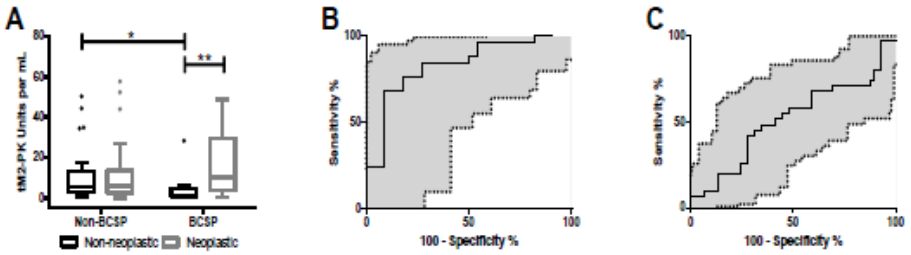



**Fig. 4**. A) Box and whisker plot demonstrating faecal tM2-PK concentration in patients recruited via BCSP or non-BCSP route, and with or without adenomatous pathology (*p<0.05, **p<0.01, by Kruskal Wallis 1-way ANOVA and Dunn's post-hoc analysis). B) ROC curve demonstrating the diagnostic accuracy of faecal tM2-PK ELISA in patients recruited following colonoscopy referral from the BCSP. C) ROC curve demonstrating the diagnostic accuracy of faecal tM2-PK ELISA in patients not recruited via the BCSP. Shaded areas represent 95%CIs.

population. The faecal tM2-PK concentration was higher in males, with age appearing to have no impact on the abundance of this enzyme. In a clinical setting, this lack of variation with age could be deemed beneficial.

Based on the manufacturer's threshold value of 4 U/mL, we found that the assay's sensitivity to detect any adenomatous pathology was 72% with specificity of 48%. This is consistent with the findings of a meta-analysis from 2012, which examined 17 studies and found the quoted sensitivities, for colorectal cancer to range from 68-97% and adenoma detection to range from 28-76%.Within the meta-analysis several of the studies focused solely on the detection of carcinoma, rather than adenomatous polyps.

To optimise the diagnostic accuracy of a faecal tM2-PK assay within a future screening programme the diagnostic threshold for the assay could be adjusted away from the manufacturer's recommended threshold of 4 U/mL. To understand whether this process may be useful, we applied a series of thresholds to our data set and calculated binomial McNemar statistics to define the threshold with optimal diagnostic utility for our cohort. Using this methodology, we demonstrated optimal utility with a threshold of 4.8 U/mL. This led to an improved sensitivity of 69.0% and specificity of 56.4% in our cohort. At this threshold 40.2% of tests yielded true positive results, 22.7% a true negative, whilst 17.5% were false positive results and 19.5% were false negative results. One of the obvious weaknesses of this methodology is that with this statistical approach the consequences of false negative and false positive results are conferred equal weighting. This is clearly not the case in this clinical situation where the consequences of a false negative result manifestly outweigh those of a false positive.

Using the manufacturer's recommended threshold of 4 U/mL, the faecal tM2-PK assay had a lower false negative rate of 16.5%, but an increased false positive rate of 20.6% in our cohort. Whilst this threshold could reduce the number of patients with adenomatous disease who were not detected by the faecal tM2-PK assay, it would lead to an increased number of endoscopic procedures that yielded no adenomatous pathology being performed. This false positive rate is substantially higher than that reported by Tonus et al. but is consistent with a study from 2006 which reported a low specificity with high false positive rate [15], suggesting that it may be within the range expected in a clinical cohort. One study from the USA estimated that a 5-10% false positive rate during colorectal screening would amount to an expense of US$6 billion if the 100% participation rate is assumed [16]. The current UK BCSP reports a sensitivity of 36.5% and specificity of 92.2% for g-FOBt. The literature suggests that if a different screening tool were to be implemented, for example i-FOBt, a higher detection rate could be achieved e.g. a sensitivity of 70.9% and specificity of 96.3% [17]. The consequence of a false negative test in bowel cancer screening is grave and therefore the lower specificity associated with tM2-PK testing will mean it is unlikely to replace either of these methods in colorectal screening or as a preliminary test in a secondary care based population. False positive results have been reported in inflammatory bowel disease, infective disease and diverticulosis [18, 19]; this is replicated in our cohort as those with non neoplastic pathology and no prior FOBt had a false positive rate of 50%.

The combination of FOBt with faecal tM2-PK ELISA has been proposed as a strategy to optimise the sensitivity and specificity for a screening programme. When we segregated our cohort into those recruited from the BCSP and those with other indications for colonoscopy, we demonstrated greater diagnostic utility for tM2-PK ELISA in patients recruited from the bowel cancer screening population (Fig. 4). This was associated with a significantly lower median tM2-PK concentration in the faeces of patients without adenomatous disease in the BCSP cohort compared to those without adenomatous disease in the non-BCSP cohort. This suggests that the utility of faecal tM2-PK as a biomarker for colonic adenomatous disease may be limited to the relatively homogeneous asymptomatic screening cohort, rather than the more heterogeneous group of symptomatic patients.

A study from 2014 assessed the utility of g-FOBt with faecal tM2-PK ELISA prior to potential colonoscopy: 1800 individuals were invited to participate, with a response rate of 54% (n=978), culminating in over 800 analysable samples being returned and 186 colonoscopies being performed. Overall, positivity was significantly increased by the addition of tM2-PK (27%). Had only i-FOBt been tested, 77% (n=189) fewer patients would have had a positive faecal test and been identified for a screening colonoscopy and 70% (n=35) fewer

patients would have had polyps detected and removed [20]. This study also reported that 10% (n=99) of tM2-PK samples could not be analysed, because of a lapse in the 48 hour window required for its processing. This is clearly an issue for a postal based screening programme as is currently in operation in the UK. A further study from Italy assessed the impact of combined testing on 280 samples. For colorectal cancer detection, i-FOBt was the test with the highest specificity and positive predictive value (0.89 and 0.53), whereas tM2-PK had the highest sensitivity and negative predictive value (0.87 and 0.96). It demonstrated a good ability of the combined test to identify colorectal cancer because in patients showing positivity to both markers, the risk of cancer was as high as 79% [21]. Other assays have been assessed in combination with FOBt, including faecal calprotectin, which demonstrated sensitivity for the detection of colorectal cancer of 79%, whereas the sensitivity for the combination of i-FOBt and tM2-PK was 93% [22].

An analysis of 697 patients in the Cleveland Clinic Foundation Adenoma Registry showed that, compared with one to two small adenomas, the risk is increased 10-fold after removal of multiple adenomas at least one of which is larger than 1 cm [23], and adenomas of >1cm diameter identify individuals at increased risk of colorectal cancer even if removed [24]. It is therefore clinically important to identify polyps of >1cm diameter. With increasing size, one might predict increased abundance of tM2-PK; however, within our cohort this was not the case.

There is a similar increase in risk for those with ≥5 individual polyps, even if they are all <1cm in size. The apparent utility of the tM2-PK assay to detect polyps irrelevant of size makes it appealing for use in a screening or surveillance programme. Another beneficial element of this finding is that polyps <1cm do not tend to bleed and, therefore, may be missed by FOBt [25]. The tM2-PK assay does not have this problem as it does not rely on the presence of blood and, therefore, has the potential to detect non-bleeding adenomatous disease. Moreover, in our cohort no difference in the abundance of tM2-PK was identified in those with five or more adenomas, compared to those with fewer.

The application of tM2-PK shows potentially beneficial outcomes when used in a homogeneous clinical group, such as the BCSP. The interference from a heterogeneous population, such as our secondary care cohort appears to impact upon the diagnostic utility of faecal tM2-PK analysis. A dedicated randomised controlled study would be required to explore this potential combination further. We must also concede that the relatively small number of patients included in this study is a limitation.

## CONCLUSIONS

In our secondary care cohort, the faecal tM2-PK assay appears to have a sensitivity that is comparable to, and in some instances, superior to existing faecal markers used as part of colorectal cancer screening programmes. The present study demonstrates a superior sensitivity to that associated with g-FOBt. However, the main issues that will prevent its independent use are its comparatively low specificity, a high false positive rate and time constraints for sample collection. Its use with existing FOBt methods does however seem appealing. Combining the two methods appears to rectify some of the flaws of each, and could improve the accuracy of an initial faecal biomarker screening tool. One role may be the sequential use of g-FOBt followed by tM2-PK. However, the introduction of this type of screening would represent additional time and financial burdens that may constrain the use of these investigations. Ultimately, in order to justify their use, a reduction in mortality from colorectal cancer with such screening tools would need to be demonstrated.

**Conflicts of interest:** No conflict of interest to declare.

**Authors' contribution**: A.B. was responsible for collecting samples, conducting laboratory work, data analysis, writing and editing the manuscript. M.B. assisted with laboratory work, data analysis, writing and editing the manuscript. D.S. collected samples and assisted in editing the manuscript. C.P. and B.C. devised the study and were responsible for writing and editing the manuscript.

**Appendix 1.** 2x2 table used for McNemar test when using a cut-off of 4.0U/mL.

| | Colonoscopy positive | Colonoscopy negative |
|---|---|---|
| Faecal tM2-PK positive | 42 | 20 |
| Faecal tM2-PK negative | 16 | 19 |

## REFERENCES


1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013; 49: 1374–1403. doi: 10.1016/j.ejca.2012.12.027
2. Rees CJ, Bevan R. The National Health Service Bowel Cancer Screening Program: the early years. Expert Rev Gastroenterol Hepatol 2013; 7: 421–437. doi: 10.1586/17474124.2013.811045
3. Trueman P, Lowson K, Chaplin S, Wright D, Duffy S, Officer I. Bowel Cancer Services : Costs and Benefits. 2007.
4. Vogelstein B. Genetic alterations during colorectal-tumor development. N Engl J Med 1988; 319: 525–532. doi: 10.1056/NEJM198809013190901
5. Network NCI. Network NCI. Colorectal Cancer Survival by Stage. 2012. Available from http://www.ncin.org.uk/publications/data_briefings/colorectal_ cancer_survival_by_stage.aspx
6. Altobelli E, Lattanzi A, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. Prev Med 2014 ; 62: 132–141. doi: 10.1016/j.ypmed.2014.02.010
7. Sawbridge D, Probert C. Population-based screening in colorectal cancer - current practice and future developments: faecal biomarkers review. J Gastrointestin Liver Dis 2014; 23: 195–202.
8. Colorectal cancer association of Canada. A guide to FOBT and FIT tests. Guaiac-based FOBT & Immunochemical-based FOBT. 2016. http://www.colorectal-cancer.ca/en/screening/fobt-and-fit/
9. Tonus C, Neupert G, Sellinger M. Colorectal cancer screening by non-invasive metabolic biomarker fecal tumor M2-PK. World J Gastroenterol 2006; 12: 7007–7011. doi: 10.3748/wjg.v12.i43.7007
10. Koss K, Maxton D, Jankowski JA. Fecal dimeric M2 pyruvate kinase in colorectal cancer and polyps correlates with tumor staging and surgical

intervention. Colorectal Dis 2008; 10: 244–248. doi: 10.1111/j.1463-1318.2007.01334.x

11. Mazurek S. Pyruvate kinase type M2: a key regulator of the metabolic budget system in tumor cells. Int J Biochem Cell Biol 2011; 43: 969–980. doi: 10.1016/j.biocel.2010.02.005
12. Tonus C, Sellinger M, Koss K, Neupert G. Faecal pyruvate kinase isoenzyme type M2 for colorectal cancer screening: a meta-analysis. World J Gastroenterol 2012; 18: 4004–4011. doi: 10.3748/wjg.v18.i30.4004
13. Kumar Y, Tapuria N, Kirmani N, Davidson BR. Tumour M2-pyruvate kinase: a gastrointestinal cancer marker. Eur J Gastroenterol Hepatol 2007; 19: 265-276. doi: 10.1097/MEG.0b013e3280102f78
14. Lüftner D, Mesterharm J, Akrivakis C, et al. Tumor type M2 pyruvate kinase expression in advanced breast cancer. Anticancer Res 2000; 20: 5077–5082.
15. Shastri YM, Naumann M, Oremek GM, et al. Prospective multicenter evaluation of fecal tumor pyruvate kinase type M2 (M2-PK) as a screening biomarker for colorectal ecoplasia. Int J Cancer 2006; 119: 2651–2656. doi: 10.1002/ijc.22243
16. Ahlquist DA. Fecal occult blood testing for colorectal cancer. Can we afford to do this? Gastroenterol Clin North Am 1997; 26: 41-55. doi: 10.1016/S0889-8553(05)70282-X
17. Shah R, Jones E, Vidart V, Kuppen PJ, Conti JA, Francis NK. Biomarkers for early detection of colorectal cancer and polyps: systematic review. Cancer Epidemiol Biomarkers Prev 2014; 23: 1712–1728. doi: 10.1158/1055-9965.EPI-14-0412
18. Shastri YM, Povse N, Schroder O, Stein J. Comparison of a novel fecal marker fecal tumor pyruvate kinase type M2 (M2-PK) with fecal calprotectin in patients with inflammatory bowel disease: a prospective study. Clin Lab 2008; 54: 389–390.
19. Chung-Faye G, Hayee B, Maestranzi S, Donaldson N, Forgacs I, Sherwood R. Fecal M2-pyruvate kinase (M2-PK): a novel marker of intestinal inflammation. Inflamm Bowel Dis 2007; 13: 1374-1378. doi: 10.1002/ibd.20214
20. Leen R, Seng-Lee C, Holleran G, O'Morain C, McNamara D. Comparison of faecal M2-PK and FIT in a population-based bowel cancer screening cohort. Eur J Gastroenterol Hepatol 2014; 26: 514-518. doi: 10.1097/MEG.0000000000000025
21. Parente F, Marino B, Ilardo A, et al. A combination of faecal tests for the detection of colon cancer: a new strategy for an appropriate selection of referrals to colonoscopy? A prospective multicentre Italian study. Eur J Gastroenterol Hepatol 2012; 24: 1145–1152. doi: 10.1097/MEG.0b013e328355cc79
22. Ardizzoia A, Parente F, Marino B, et al. A combination of fecal tests for the detection of colon cancer: A new strategy for appropriate prioritization of referrals to colonscopy- A prospective Italian study. J Clin Oncol 2011; 29:(suppl; abstr 1580).
23. Noshirwani KC, van Stolk RU, Rybicki LA, Beck GJ. Adenoma size and number are predictive of adenoma recurrence: implications for surveillance colonoscopy. Gastrointest Endosc 2000; 51: 433–437. doi: 10.1016/S0016-5107(00)70444-5
24. Atkin WS, Saunders BP; British Society for Gastroenterology; Association of Coloproctology for Great Britain and Ireland. Surveillance guidelines after removal of colorectal adenomatous polyps. Gut 2002; 51 Suppl 5: V6–V9. doi: 10.1136/gut.51.suppl_5.v6
25. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 2008; 134: 1570–1595. doi: 10.1053/j.gastro.2008.02.002

Curr Colorectal Cancer Rep
DOI 10.1007/s11888-016-0308-7



# How Can We Improve Adenoma Detection Rate?

**Ashley Bond** [1] · **Sanchoy Sarkar** [1]



**Abstract** Adenoma detection rate (ADR) is a considerable component of colonoscopy quality assurance; it has a clear link to future morbidity and mortality from colorectal cancer. There are a number of potential factors, both modifiable and non-modifiable that can impact upon ADR. Modes of improving ADR include techniques and behaviours during colonoscopy, image enhancement techniques, technological advancements, advancements in endoscope designs, developments in accessories and continued medical education. However, whilst a number of technologies are emerging to improve ADR, at present, it seems that education, team work and optimising current practice will provide the biggest gains in ADR whilst maintaining financial acceptability. The review aims to present a balanced summary of the evidence currently available and does not propose to serve as a guideline.

**Keywords** Colonoscopy · Adenoma · Detection rate · Colorectal cancer · Screening

## Introduction

Colorectal cancer is a considerable worldwide health issue, being the third most common cancer in men and the second in women. Worldwide, an estimated 1.2 million cases of colorectal cancer occur annually [1]. The highest incidences have previously been found in Europe, New Zealand, Japan and the USA. However, more recently, these levels have begun to plateau and this is thought to be due to the relatively recent implementation of national bowel cancer screening programmes which has shown to reduce the risk of death from colorectal cancer through detection of tumours at an earlier, at a more treatable stage and through removal of precancerous adenomas [2].

There are a number of quality assurance measures inherent to these screening programmes, these include caecal intubation rate, bowel preparation quality, complications, cancer detection and adenoma detection rate (the proportion of colonoscopies performed by a physician that detect at least one histologically confirmed colorectal adenoma). Following two recent landmarks, adenoma detection rate (ADR) is now considered to be the most important quality assurance measure within screening colonoscopy. The first, demonstrated increased risk of interval cancer when the colonoscopy is performed by an endoscopist with an ADR below 20 % [3••]. As a result, professional societies recommend a detection rate of >25 % in order to be deemed adequate [4]. The second supported this inverse relationship of ADR attainment and improved colorectal cancer outcomes, with each 1.0 % increase in ADR was associated with a 3.0 % reduction in the risk of interval cancers and a 5 % reduction in CRC mortality [2]. A recent microsimulation modelling study further emphasised this by demonstrating that a higher adenoma detection rate in screening colonoscopy was associated with a lower lifetime risk of colorectal cancer and colorectal cancer mortality without being associated with higher overall costs [5].

With this in mind, the question of how we can improve ADR is a pertinent one, as we strive to improve outcomes from national screening programme and overall cancer-related mortality. This article will discuss how ADR can be improved via optimisation of current practice, the addition of

Topical Collection on *Prevention and Early Detection*

✉ Sanchoy Sarkar
sanchoy@aol.com

[1] Department of Gastroenterology and Hepatology, Royal Liverpool and Broadgreen University Hospital Trust, Prescot Street, Liverpool, UK

Published online: 20 January 2016

new behaviours and techniques to current practice, the importance of continued medical education, the potential impact of new technologies and imaging techniques.

## Optimisation of Current Practice

### *Bowel Preparation*

Poor bowel preparation has been associated with an adenoma miss rate of 43 % [6]. Consequently, the American Consensus Guidelines strongly recommended that optimal split-dose bowel preparation should be employed in order to improve ADR [7]. Studies have demonstrated a substantial improvement in ADR (35 %) with split-dose preparation ($p \leq 0.001$). They also showed a clear improvement in other quality measures such as caecal intubation rate (95.5 %) and overall bowel preparation quality [8]. All attempts should be made therefore to optimise compliance with the prescribed split-dose regime. This can include pre-procedural counselling, either face-to-face or telephone. One such study assessed the impact of telephone education relating to the bowel preparation prior to colonoscopy. There was an improvement in the aforementioned quality assurance measures, including ADR [9]. Pre-procedural counselling, including bowel preparation educational, forms part of the UK BCSP.

### *Use of Antispasmodics and Sedation*

There are conflicting results in a number of studies relating to the use of antispasmodics and ADR. A recent meta-analysis based upon the finding of eight RCTs conducted in Europe, Asia and Australia concluded that hyoscine use in patients undergoing colonoscopy does not appear to significantly increase the detection of adenomas [10]. Contrary to this, a study examining the use of different antispasmodic that is applied topically during the procedure: L-menthol (an organic compound found in peppermint oil), demonstrated an improved ADR [11]. A further study that demonstrated improved ADR with the use of hyoscine was conducted within the UK BCSP, specifically there was an increase in the number of polyps, the detection of those greater than 1 cm and all right-sided polyps ($p \leq 0.001$) [12•].

The decision to use sedation is often a one based on clinical parameters and patient choice. The evidence appears to suggest as a depth of sedation increases so does the ADR but at the expense of an increasing complication rate. This is evidenced when you consider the results of a study that demonstrated a larger amount of mono or dual (conscious) sedation improved many aspects of colonoscopy assurance measures. ADR increased (25.9 to 35 %), early complications rate decreased (3.4 to 0.3 %) and completion rates increased (88.3 to 96.4 %) [13]. When higher levels of sedation are used (using propofol), studies have suggested an increase in ADR; however, these also showed an increase in complication rates, particularly perforation [14]. Conversely, a study of 3252 colonoscopies showed no difference for those receiving propofol and conscious sedation (midazolam and fentanyl) [15]. Current practice of conscious sedation appears to optimise ADR whilst minimising potential complications and therefore if permitted by patient choice and clinical factors should be employed.

### *Insertion and Withdrawal Times*

Conventionally, visualisation and polypectomy are performed on withdrawal. However, studies have suggested that this may not be optimal for small polyps, with one such study finding that polyp <10 mm identified during insertion can be easily missed on withdrawal as the colonic configuration changes, suggesting polypectomy during insertion is better [16]. A recent study within the Bowel Cancer Screening Programme (BCSP) in England demonstrated a plateau effect after a withdrawal time of approximately 10 min. The lowest ADR was demonstrated if the withdrawal was less than 7 min, with the maximum ADR seen with a withdrawal time of 9–11 min [12•]. A separate study concluded a mean withdrawal time of >8 min was the only modifiable factor to impact upon ADR in colorectal cancer screening colonoscopies [17••]. A study from 2014 compared 610 colonoscopies where patients were randomised to either polypectomy during insertion and withdrawal or just withdrawal. This study did not demonstrate any difference in ADR between the two groups [18]. Overall, it would appear that ADR can be improved by adopting a withdrawal time of between 8 to 11 min and performing polypectomy of small polyps during the insertion phase.

### *Other Procedural Factors*

There are other potential elements of colonoscopy that could enhance ADR. One such factor is dynamic position change. Randomised controlled trials exploring this factor have produced conflicting results, but overall appear to show predominately positive findings. Position change should form part of the routine procedure as it allows for improved negotiation of the colon thereby improving caecal intubation rate and patient comfort. The impact of dynamic position changes during colonoscopy result in efficient colonic distension with less insufflation of air, improved mucosal visualisation via the shifting of fluids and residues and opening tight angles at flexures [19]. Dynamic position change therefore has the potential to improve numerous quality assurance measures and in a number of studies has repeatedly been shown to improve ADR [20, 21]. As a result of this, in the UK, dynamic position change is taught as standard practice to trainee colonoscopists. The start time of the procedure can also be modified in order to optimise ADR. Procedures starting in the second half of a

session (11:00–14:00 or 16:00–18:00) have been associated with a reduction in detection of adenomas and advanced adenomas compared with procedures starting between 08:00 and 11:00 or 14:00 and 16:00 [12•]. Clearly though in a health care system such as the NHS, capacity and efficiency need to be optimised, meaning all available procedural time should be utilised. An "all eyes on screen" technique was devised as part of one study. In doing so, all members of the team were asked to observe the screen and spot adenomas as the procedure was in progress; this was shown to improve ADR. The study in question demonstrated an increase in ADR from 34 to 51 % in 2 years [22]. Central visual gaze of the colonoscopist has also been shown to increase ADR [23]. Clearly therefore, it is important to prevent distractions within the room during colonoscopy.

*Addition of New Behaviours and Techniques to Current Practice*

The addition of water infusion techniques have been suggested as a method of improving colonoscopy quality measures, including ADR. It has been shown to improve caecal intubation, reduce colonic spasms, lower patient discomfort and need for sedation [24, 25]. Overall, studies have failed to show this improvement when considering ADR. A study in 2009 assessed the potential benefit of this technique as a method of improving ADR; it did not show any benefit, failing to reach statistical significance.[26]. A systematic review performed in 2012 comparing water infusion to conventional insufflations reported no difference in ADR [27]. Together with the lack of evidence suggesting an inferred benefit on ADR, the prolonged insertion time, colonoscopist experience and general technicalities of these techniques are unlikely to become common place in routine practice. A second intra-procedural technique that has the potential to improve ADR is caecal retroflexion. The impact of such was examined by a prospective cohort study conducted in the USA. During this study, 1000 consecutive adults undergoing colonoscopy were included. A standard forward-viewing colonoscopy was performed and polyps were removed. A repeated examination in retroflexion of the caecum to the hepatic flexure was then performed. Then, 94.4 % of the patients had successful retroflexion in the caecum. This second examination in retroflexion demonstrated a 9.3 % miss rate for the forward-viewing method [28]. Risk of complication, namely perforation, is likely to limit the use of this technique.

*Continued Medical Education and Quality Improvement Programmes*

The potential impact of the individual colonoscopist upon ADR, i.e., the person performing the examination, is well described by the literature; an example of which can be seen via a study that examined 12,000 screening colonoscopy in order to assess factors that impact upon quality assurance measures. This study found that annual case volume and life experience did not affect ADR, with continued medical education (CME) being the most influential factor on ADR, specifically describing a higher ADR in those who attended most CME meetings [29]. Formal assessment of the impact of a CME programme was performed by the Mayo clinic. For this study, an additional training program, known as Endoscopic Quality Improvement Program (EQUIP), was drawn upon. Baseline ADRs were recorded, followed by half of the 15 colonoscopists being randomly assigned to the EQUIP. A comparison was then made between the baseline ADRs and those seen after the implementation of the EQUIP. There were a total of 1200 procedures completed in each of the two study phases. There was a significant impact upon ADR from the EQUIP. The group of endoscopists randomized to EQUIP training had an overall increase in ADR to 47 %, whereas the ADR for the endoscopists not in the EQUIP saw their ADR unchanged at 35 % [30••, 31]. All the evidence supports CME as one of the most effective methods of improving and maintaining superior ADRs. A further study examining factors that influenced the quality of colonoscopy again identified CME as a key factor. This study looked at 12,134 colonoscopies in Germany over an 18-month period. The influencing factors that were related to the colonoscopists were the number of CME meetings attended ($p = 0.012$) and instrument generation ($p = 0.001$) [29].

The BCSP in England have opted for third party reporting and quality assurance of colonoscopy with continuous audit feedback of individual performance parameters. This has led to improvement in a number of quality assurance measures, including ADR and caecal intubation rates. Such an improvement can be seen in the ADR of an author of this paper where an already experienced colonoscopist can year upon year improve their performance (Table 1).

## New Technology

*Endoscope Design and Adjuncts*

There are a number of available and emerging technologies that have been reported to improve ADR. One such technology is extra-wide-angle-view colonoscopes. The Full Spectrum Endoscopy (FUSE) system (EndoChoice) is currently available on the market, but not in routine use. It allows for a 330° view of the colonic lumen, thereby increasing visualisation of the mucosa when compared to a standard forward-viewing scope (Fig. 1). In addition to the standard forward-viewing camera, on the left and right side of the FUSE scope's tip, there are two additional cameras. This results in the transmitted images from the cameras being displayed on three adjoining monitors corresponding to each camera. This

**Table 1** Table demonstrating the cumulative improvement in an individual ADR by third party reporting and quality assurance within the BCSP in England

| Year | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---|---|---|---|---|---|---|---|
| ADR (%) | 39.34 | 36.96 | 50 | 42.96 | 50.98 | 54.41 | 62.04 |

configuration is stated to provide an enhanced view of the colonic lumen, including visualisation of the blind spots at the flexures or proximal edges of the mucosal folds [31].

Studies using both anatomical models and patients undergoing colonoscopy have shown an improvement in ADR when compared to a conventional forward-viewing scope. For example, a study using 37 colonoscopies carried out on anatomical models using the forward-viewing camera scope, followed by a colonoscopy with full FUSE system active, i.e., all three cameras on saw an increase in ADR. A total of 85.7 % of the polyps placed within the model were detected with the three cameras whereby only 52.9 % were detected with only the forward-viewing screen active ($p \leq 0.001$). Some of the polyps were placed in notoriously difficult positions, i.e., behind fold; there was a significant increase in detection of such polyps using the FUSE system (81.9 vs. 31.9 %) [32]. The application of FUSE and its impact upon ADR was again assessed in 2014. In this international, multicentre, randomised trial, 185 participants were assessed. The adenoma miss rate was significantly lower in patients in the FUSE group than in those in the standard forward-viewing procedure group: 7 vs. 41 % ($p = 0.0001$) [33]. Larger studies will need to be conducted in order to assess this improvement further, but all the evidence appears to infer a significant improvement in ADR.

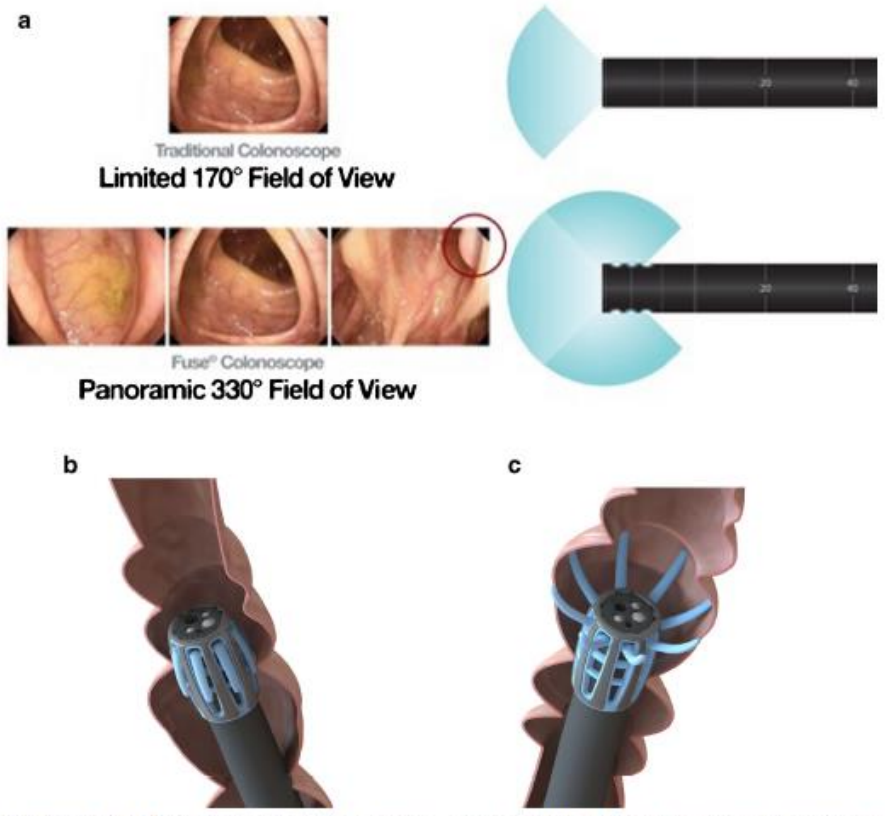



**Fig. 1** *Top*: image representing the viewing screen of the FUSE system and the potential increased field of view (images courtesy of EndoChoice, Inc., Alpharetta, GA). *Bottom*: image demonstrating the Endocuff with the wings collapsed during insertion and splayed during withdrawal (images courtesy of Arc Medical Design Ltd., Leeds, England, United Kingdom)

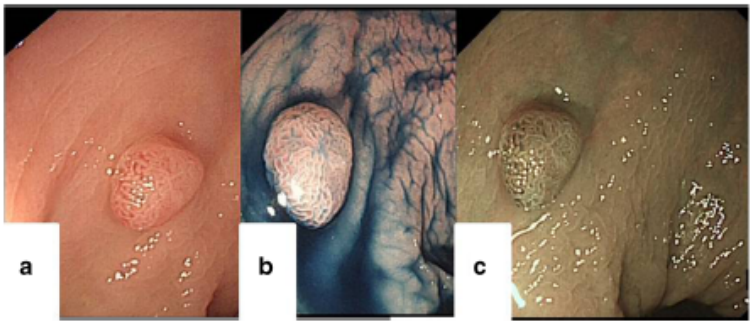



**Fig. 2** **a** Adenomatous polyp viewed in normal white light. **b** The same polyp following colonic dye spray with indigo carmine. **c** Polyp viewed with narrow band active

The NaviAid G-EYE colonoscope (SMART Medical Systems) is another new system that has been reported to improve ADR. It is an example of balloon-assisted colonoscopy; for this system, there is an integrated balloon on the flexible tip of the scope. The balloon is inflated and deflated during the withdrawal process. The mechanical flattening and straightening of haustral folds with the inflated balloon permit visualisation of hidden anatomic areas, thus increasing the ADR [25]. The only studies that exist for this system are based on anatomical models. One such study showed a significantly greater ADR in the balloon-assisted group ($p \leq 0.0001$) [34]. Early signs are promising, but larger-scale clinical trials will be required.

Recent developments in endoscope adjuncts have provided another potential avenue to improve ADR. One such adjunct is Third Eye Retroscope® (Avantis Medical Systems, Inc). The system consists of a video processor, a single-use polarizing filter cap for the colonoscope light source, and a 3.5-mm flexible single-use catheter with a camera and diode light source at the tip. The retroscope is retroflexed 180° after being advanced through the working channel of the colonoscope and provides a 135° retrograde view of the colon [24, 31]. Anatomical model studies showed that standard colonoscopy detected 12 % of the polyps located on the proximal aspects of folds, whilst 81 % of these polyps were detected with the Third Eye Retroscope [35]. TERRACE is the largest study examining the use of Third Eye Retroscope. It was a multicentred study that included 349 patients. It demonstrated an increase ADR of 23.2 %; there was an average increase in withdrawal time of 2 min to achieve this [36]. The use of this device means that the working channel is occupied and therefore impairs suction, adding a further complexity to its use. Avantis Medical Systems also produce third eye panoramic. The device clips onto the tip of a standard colonoscope to provide two additional video cameras. Combining the three images results in a 330° "panoramic" view, much like the FUSE system. Early studies of this device have demonstrated an ADR of 44 % [37].

Cap-assisted colonoscopy is another adaptation to conventional colonoscopy that has been reported to improve ADR. Endocuff (Arc Medical, USA), currently available on the market, is one such device (Fig. 1). The Endocuff (EC) is a 2-cm-long flexible cuff with two rows of small flexible, hinged wings that help to flatten large mucosal folds during withdrawal of the instrument but lay flush to the scope during insertion [31]. A study examining the impact of Endocuff looked at 498 patients undergoing screening colonoscopy. It showed that the addition of the device increased the absolute rate of polyp detection by 14 % when compared to unassisted colonoscopy, with absolute ADR increasing from 42 to 56 % ($p = 0.001$). This improvement was particularly noted in the sigmoid colon [38]. A third cap type device is the EndoRings (EndoRings; Endo-Aid, Caesarea, Israel). This is a single-use device composed of a series of three clear silicone discs or rings positioned sequentially on a cylindrical cuff. These rings engage and mechanically stretch the mucosa and colon folds during colonoscopy. The rings also provide some traction to maintain position during loop reduction, to decrease slippage and to maintain stability during instrumentation [24]. The use of this device was assessed in the CLEVER study. This study concluded that colonoscopy with EndoRings has lower adenoma and polyp miss rates than standard colonoscopy, which may improve the efficacy particularly of screening and surveillance colonoscopies [39].

Other cap devices are available; these are mainly transparent tips attached to the distal end of the colonoscope. These were initially developed to aid endoscopic mucosal resection, but have been proposed as a method of improving ADR. This has been shown to be true in the hands of trainees and less-experienced colonoscopist [40]. More recently, a randomised trial that included 1113 patients examined the impact of cap-assisted colonoscopy on ADR, demonstrated a 20 % increase in ADR for some endoscopists, whilst others saw a 15 % decrease. Cap-assisted colonoscopy was associated with a shorter caecal intubation time and increase TI intubation rate across all colonoscopists. The authors subsequently

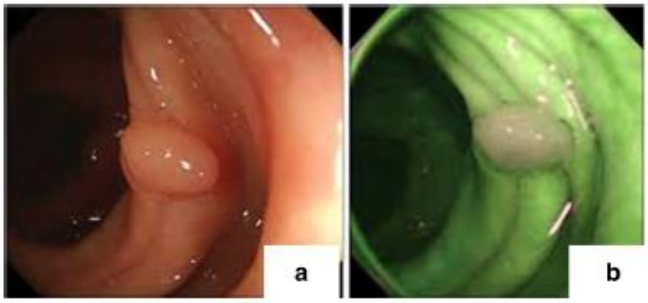



**Fig. 3** **a** Adenomatous polyp viewed in white light. **b** The same polyp viewed with digital autofluorescence (AFI)

concluded cap-assisted colonoscopy may be beneficial for selected endoscopists [41].

*Enhanced Imaging Techniques*

There is evidence to support the use of high-definition systems in order to improve ADR. In one study, this was particularly so when advances systems were used by colonoscopists with ADR <20 %. However, for those with an ADR already >20 %, there was no improvement in detection of high-risk polyps, flat polyps or proximal lesions [42]. A meta-analysis involving 4422 patients concluded that high-definition systems are superior in the detection of smaller polyps, without improving the detection of larger and higher risk lesions [43]. A further study with similar design also showed a lower adenoma miss rate with high-definition colonoscopy [44]. The available evidence supports the use of high-definition systems as a means to improve ADR and highlights it as a factor that can improve overall quality without detrimental effects on other measures [29].

Chromoendoscopy is another method by which ADR can potentially be improved. The impact of chromoendoscopy on ADR is described in a number of studies, many of which have demonstrated an increase in the yield of neoplasia detection [45–47]. Chromoendoscopy is recommended for screening in high-risk groups, such as IBD, and is particularly useful in detecting flat lesions. A study of 660 average risk patients compared chromoendoscopy to standard white light and did not demonstrate any significant difference in ADR [48]. Given the increased procedural time, it is likely that chromoendoscopy will remain in place for those undergoing colonoscopy for surveillance of high-risk conditions. Newer stains are receiving interest as they may remove the time constraints of conventional chromoendoscopy. One such formulation uses methylene blue (MB MMX, Cosmos Technologies). This has been designed as a modified release device which ensures colonic release. The methylene blue is not taken up by the adenomatous tissue but is by the adjacent normal mucosa, thus providing the contrast required for visualisation [31]. A study of this compound demonstrated an ADR of 63.5 % when used in 96 patients undergoing routine colonoscopy [49].

**Table 2** Summary table of techniques and technologies that may impact upon ADR. Includes source of evidence and potential financial cost impact of its use/implementation

| Technology or technique | Improves ADR | Evidence | Cost impact |
|---|---|---|---|
| Bowel preparation | Yes | RCT's | + |
| Withdrawal time | Yes | Observational studies | + |
| Use of sedation | Mixed | RCT's | + |
| Use of antispasmodics | Mixed | RCT and meta-analysis | + |
| Dynamic position change | Yes | RCT's | + |
| "All eyes on screen" | Yes | RCT's | + |
| Continued medical education | Yes | RCT's | ++ |
| Narrow band imaging | No | RCT's | +++ |
| High-definition systems | Yes | RCT's | +++ |
| Full spectrum endoscopy (FUSE) | Yes | RCT's (overall more evidence needed) | +++ |
| Balloon-assisted devices | Too early | Anatomical models only | +++ |
| Endocap | Yes | RCT'S | ++ |
| Endocuff | Yes | RCT'S | ++ |
| EndoRing | Possibly | RCT's | ++ |

Digital chromoendoscopy can further enhance the quality of colonic dye spray. Narrow band imaging (NBI), available on the Olympus systems, is one such technology (Fig. 2). Fujinon Intelligent colour Enhancement (FICE) and the Pentax technology equivalent, i-Scan, are examples of other systems that narrow the bandwidth of conventional white-light colonoscopy to improve visualisation. There is, however, a lack of evidence to support their use in order to improve ADR, with many concluding NBI does not increase the yield of colon polyps, adenomas, or flat adenomas, nor does it decrease the miss rate of colon polyps or adenomas in patients undergoing screening/surveillance colonoscopy [50–52]. The use of these systems is best placed in assessment of polyps once they have been detected.

Digital autofluorescence (AFI) appears to be an enhanced imaging technique that can improve ADR (Fig. 3). The use of AFI has been shown to improve ADR in both routine and high-risk colonoscopy. A study examining AFI use in those undergoing colonoscopy for surveillance of Lynch syndrome or a strong family history of colorectal cancer reported a significantly higher sensitivity of AFI compared with white light (92 vs. 68 %; $p=0.001$). Together with an increased ability to detect smaller adenomas when compared to standard white light alone [53]. When used by less-experienced colonoscopists, there is a significant increase in ADR and fall in miss rate [54]. Direct comparison of standard white light, NBI and AFI demonstrated a superior diagnostic accuracy for adenomatous disease in the AFI group (white light, $p=0.001$; with NBI, $p=0.016$) [55].

## Conclusion

There are a number of technological advancements both in terms of imaging, scope design and adjuncts that are proposed to improve ADR which have been summarised in Table 2. However, the largest gains in ADR can be made through optimisation of current practice, the use of high-definition colonoscopes and continued medical education. Such measures are particularly acceptable given the additional financial implications and prolonged procedural times related to many of the technological advancements. Some of these new technologies may have a role in training or at the other end of the spectrum, being used in very experienced hands for further assessment or intervention. Optimisation of bowel preparation compliance, intra-procedural position changes, observation of the screen display by all members of the team and ongoing education of the colonoscopist with quality improvement programmes can significantly improve ADR and therefore reduce future mortality from colorectal cancer. These are factors which are potentially modifiable by all endoscopy departments.

### Compliance with Ethical Standards

**Conflict of Interest** Ashley Bond and Sanchoy Sarkar declare that they have no conflict of interest.

**Human and Animal Rights and Informed Consent** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

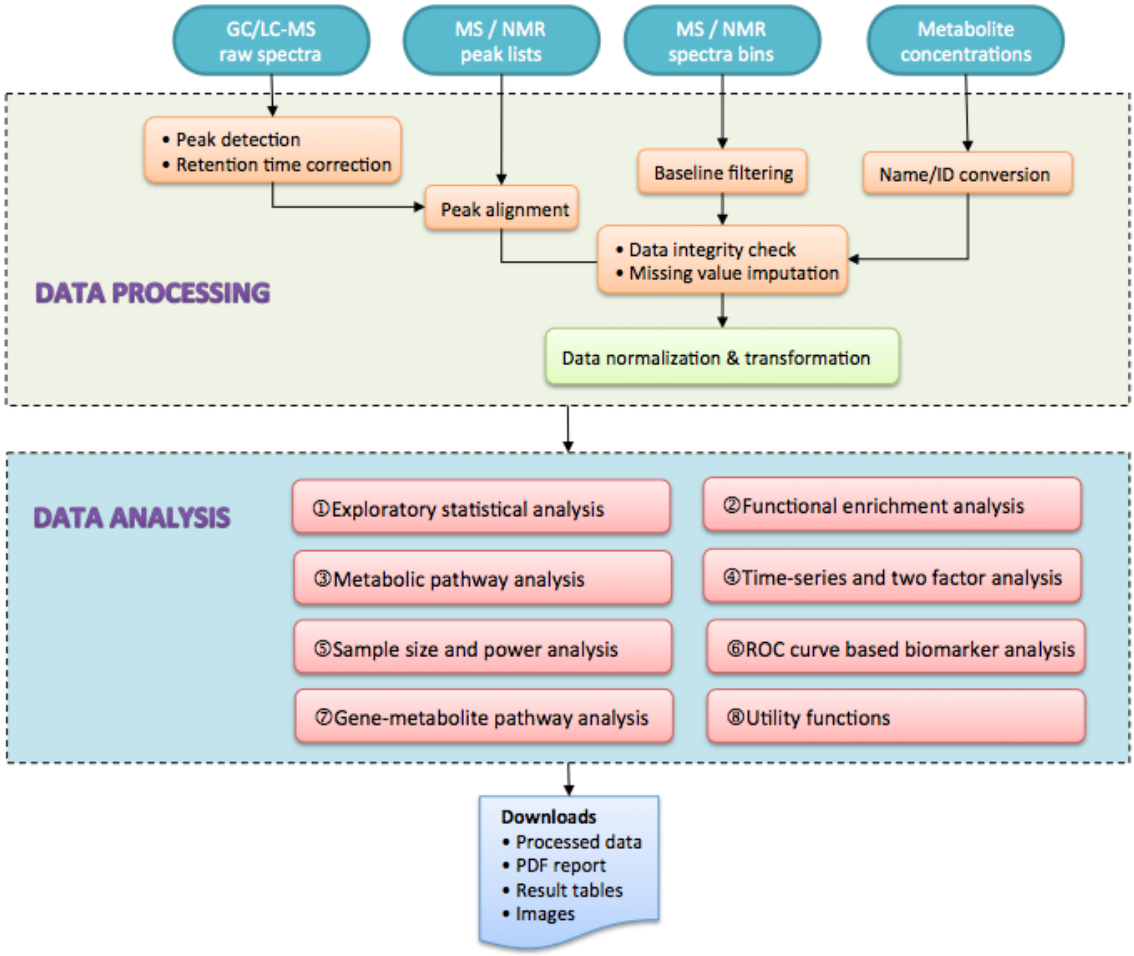

1. Ferlay J, Shin HR BF et al. GLOBOCAN 2008 v2.0. Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 10 2010. 2008.
2. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med.[Internet]. 2014;370:1298–306. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24693890.
3. •• Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362:1795–803. **This study clearly defined the importance of improving adenoma detection rates, by showing a fall in future mortality and morbidity from colorectal cancer with increasing adenoma detection.**
4. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. Am J Gastroenterol. 2006;101:873–85.
5. Meester RGS, Doubeni CA, Lansdorp-Vogelaar I, Jensen CD, van der Meulen MP, Levin TR, et al. Variation in adenoma detection rate and the lifetime benefits and cost of colorectal cancer screening. Jama [Internet]. 2015;313:2349. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2015.6251.
6. Benjamin L, Fay K, Michael Glick AJ, Rosenbaum BA, Timothy W, Alfred I, et al. The impact of suboptimal preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. Gastrointest Endosc. 2012;29:997–1003.
7. Benjamin L, Fay K, Michael G, Adam J, Rosenbaum BA, Timothy W, et al. The impact of suboptimal preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. Gastrointest Endosc. 2012;29:997–1003.
8. Gurudu SR, Ramirez FC, Edwyn H, Leighton JA, Crowell MD. Increased adenoma detection rate with system-wide implementation of a split-dose preparation for colonoscopy. Gastrointest Endosc. 2011. doi:10.1016/j.gie.2012.04.456.
9. Liu X, Luo H, Zhang L, Leung FW, Liu Z, Wang X, et al. Telephone-based re-education on the day before colonoscopy improves the quality of bowel preparation and the polyp detection rate: a prospective, colonoscopist-blinded, randomised, controlled study. Gut [Internet]. 2013;1–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23503044.

10. Ashraf I. Hyoscine for polyp detection during colonoscopy: a meta-analysis and systematic review. World J Gastrointest Endosc [Internet]. 2014;6:549. Available from: http://www.wjgnet.com/1948-5190/full/v6/i11/549.htm.
11. Inoue K, Dohi O, Gen Y, Masayasu JO, Mazaki T, Tokita K, et al. L-menthol improves adenoma detection rate during colonoscopy: a randomized trial. Endoscopy. 2014;46(03):196–202.
12.• Thomas J, Lee W, Rees CJ, Blanks RG, Moss SM, Claire N, et al. Colonoscopic factors associated with adenoma detection in a national colorectal cancer screening program. Endoscopy. 2014;46:203–11. **An excellent summary of the potentially modifiable variables that can improve adenma detection rate. This is important to any endoscopy department as it provides methods to achieve incremental gains.**
13. Triantafyllou K, Sioulas AD, Kalli T, Misailidis N, Polymeros D, Papanikolaou IS, et al. Optimized sedation improves colonoscopy quality long-term. Gastroenterol Res Pract [Internet]. Hindawi Publishing Corporation; 2015;2015:1–9. Available from: http://www.hindawi.com/journals/grp/2015/195093/.
14. Parra V, Watanabe J, Nago A, Astete M, Rodríguez C, Valladares G, et al. Sedation used during colonoscopies at the Peruvian-Japanese policlinic. Their relationship with quality indicators. Rev Gastroenterol del Perú. 2010;30:40–5.
15. Metwally M, Agresti N, Hale WB, Ciofoaia V, O'Connor R, Fine J, et al. Conscious or unconscious: the impact of sedation choice on colon adenoma detection. World J Gastroenterol. 2011;17:3912–5.
16. Wildi SM, Schoepfer AM, Vavricka SR, Fruehauf H, Safroneeva E, Wiegand N, et al. Colorectal polypectomy during insertion and withdrawal or only during withdrawal? a randomized controlled trial. Endoscopy. 2012;44:1019–23.
17.•• Jover R, Zapater P, Polanía E, Bujanda L, Lanas A, Hermo JA, et al. Modifiable endoscopic factors that influence the adenoma detection rate in colorectal cancer screening colonoscopies. Gastrointest Endosc [Internet]. Elsevier Inc.; 2013;77:381–9.e1. Available from: http://dx.doi.org/10.1016/j.gie.2012.09.027. **An excellent summary of the potentially modifiable variables that can improve adenma detection rate. This is important to any endoscopy department as it provides methods to achieve incremental gains.**
18. Sanaka MR, Parsi MA, Burke CA, Barnes D, Church J, Rizk M, et al. Adenoma detection at colonoscopy by polypectomy in withdrawal only versus both insertion and withdrawal: a randomized controlled trial. Surg Endosc [Internet]. 2014;29:692–9. Available from: http://link.springer.com/10.1007/s00464-014-3723-3.
19. Ghosh S. Dynamic position change at colonoscopy improves adenoma detection. Can J Gastroenterol. 2013;27:508.
20. East JE, Bassett P, Arebi N, Thomas-Gibson S, Guenther TSB. Dynamic position changes during colonoscope withdrawal increase adenoma detection: a randomized, crossover trial. Gastrointest Endosc. 2011;73:456–63.
21. Köksal AŞ, Kalkan İHTS. A simple method to improve adenoma detection rate during colonoscopy: altering patient position. Can J Gastroenterol. 2013;27:509–12.
22. Saunders F, Allegretti PSD. Adenoma detection rate (ADR) as a quality standard in a privately operated community surgical center. Am J Gastroenterol. 2012;107:S796–7.
23. Almansa C, Shahid MW, Heckman MG, Preissler S, Wallace MB. Association between visual gaze patterns and adenoma detection rate during colonoscopy: a preliminary investigation. Am J Gastroenterol [Internet]. Nature Publishing Group; 2011;106:1070–4. Available from: http://dx.doi.org/10.1038/ajg.2011.26.
24. Dik VK, Moons LMG, Siersema PD. Endoscopic innovations to increase the adenoma detection rate during colonoscopy. World J Gastroenterol. 2014;20:2200–11.
25. Titi M, Gupta GSP. Advanced colonoscopic imaging: do new technologies improve adenoma detection? Gastroenterol Endosc News. 2014;12.
26. Leung FW, Aharonian HS, Leung JW, Guth PHJGZ. Impact of a novel water method on scheduled unsedated colonoscopy in U.S. veterans. Gastrointest Endosc. 2009;69:546–50.
27. Leung FW, Amato A, Ell C, Friedland S, Harker JO, Hsieh YH, et al. Water-aided colonoscopy: a systematic review. Gastrointest Endosc. 2012;76:657–66.
28. Hewett DG, Rex DK. Miss rate of right-sided colon examination during colonoscopy defined by retroflexion: an observational study. Gastrointest Endosc [Internet]. Elsevier Inc.; 2011;74:246–52. Available from: http://dx.doi.org/10.1016/j.gie.2011.04.005.
29. Adler A, Wegscheider K, Lieberman D, Aminalai A, Aschenbeck J, Drossel R, et al. Factors determining the quality of screening colonoscopy: a prospective study on adenoma detection rates, from 12 134 examinations (Berlin colonoscopy project 3, BECOP-3). Gut. 2012;236–41.
30.•• Coe SG, Crook JE, Diehl NN, Wallace MB. An endoscopic quality improvement program improves detection of colorectal adenomas. Am J Gastroenterol. 2013;108:219–26. quiz 227. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23295274. **This study clearly demonstrates how an individual unit can use continued medical education to improve adenoma detection rate. Continued medical education is available to all endoscopists and therefore is a pertient topic.**
31. Bond A, Sarkar S, Bond A, Sarkar S. New technologies and techniques to improve adenoma detection in colonoscopy. 2015;7:969–80.
32. Gralnek IM, Carr-Locke DL, Segol O, Halpern Z, Siersema PD, Sloyer A, et al. Comparison of standard forward-viewing mode versus ultrawide-viewing mode of a novel colonoscopy platform: a prospective, multicenter study in the detection of simulated polyps in an in vitro colon model (with video). Gastrointest Endosc. 2013;77:472–9.
33. Gralnek IM, Siersema PD, Halpern Z, Segol O, Melhem A, Suissa A, et al. Standard forward-viewing colonoscopy versus full-spectrum endoscopy: an international, multicentre, randomised, tandem colonoscopy trial. Lancet Oncol. 2014;15:353–60.
34. Hasan N, Gross SA, Gralnek IM, et al. A novel balloon colonoscope detects significantly more simulated polyps than a standard colonoscope in a colon model. Gastrointest Endosc. 2014.
35. Triadafilopoulos G, Watts HD, Higgins JVDJ. A novel retrograde-viewing auxiliary imaging device (Third Eye Retroscope) improves the detection of simulated polyps in anatomic models of the colon. Gastrointest Endosc. 2007;65:139–44.
36. Leufkens AM, DeMarco DC, Siersema PD, et al. Effect of a retrograde-viewing device on adenoma detection rate during colonoscopy: the TERRACE study. Gastrointest Endosc. 2011;73:480–9.
37. Rubin M. Expanding the view of a standard colonoscope with the Third Eye ® Panoramic ™ cap. World J Gastroenterol. [Internet]. 2015;21:10683. Available from: http://www.wjgnet.com/1007-9327/full/v21/i37/10683.htm.
38. Biecker E, Floer M, Heinecke A, et al. Novel Endocuff-assisted colonoscopy significantly increases the polyp detection rate: a randomized controlled trial. J Clin Gastroenterol. 2014.
39. Dik VK, Gralnek IM, Segol O, Suissa A, Belderbos TD, Moons LM, et al. Multicenter, randomized, tandem evaluation of EndoRings colonoscopy - results of the CLEVER study. Endoscopy. 2015;published ahead of print.
40. Kondo S, Yamaji Y, Watabe H, Yamada A, Sugimoto T, Ohta M, et al. A randomized controlled trial evaluating the usefulness of a transparent hood attached to the tip of the colonoscope. Am J Gastroenterol. 2007;102:75–81.

41. Heiko P, Bensen SP, Arifa T, Gordon SR, Campbell Levy L, Brian B, et al. Cap-assisted colonoscopy and detection of adenomatous polyps (CAP) study: a randomized trial. Endoscopy. 2015;47:891–7.
42. Waldmann E, Britto-Arias M, Gessl I, Heinze G, Salzl P, Sallinger D, et al. Endoscopists with low adenoma detection rates benefit from high-definition endoscopy. Surg Endosc. 2015;29:466–73.
43. Subramanian V, Mannath CJ, Hawkey KR. High definition colonoscopy vs. standard video endoscopy for the detection of colonic polyps: a meta-analysis. Endoscopy. 2011;43(6):499–505.
44. Gross SA, Buchner AM, Crook JE, Cangemi JR, Picco MF, Wolfsen HC, et al. A comparison of high definition-image enhanced colonoscopy and standard white-light colonoscopy for colorectal polyp detection. Endoscopy. 2011;43:1045–51.
45. Fujii T, Rembacken BJDM, et al. Flat adenomas in the United Kingdom: are treatable cancers being missed? Endoscopy. 1998;30:437–43.
46. Kiesslich R, von Bergh MHM, et al. Chromoendoscopy with indigocarmine improves the detection of adenomatous and nonadenomatous lesions in the colon. Endoscopy. 2001;31:1001–6.
47. Jaramillo E, Watanabe MSP, et al. Flat neoplastic lesions of the colon and rectum detected by high-resolution video endoscopy and chromoscopy. Gastrointest Endosc. 1995;42:114–22.
48. Kahi CJ, Anderson JC, Waxman I, Kessler WR, Imperiale TF, Li XRD. High-definition chromocolonoscopy vs. high-definition white light colonoscopy for average-risk colorectal cancer screening. Am J Gastroenterol. 2010;105(6):1301–7.
49. Repici A, Di Stefano AF, Radicioni MM, et al. Methylene blue MMX tablets for chromoendoscopy. Safety tolerability and bioavailability in healthy volunteers. Contemp Clin Trials. 2012;33:260–7.
50. Pasha SF, Leighton JA, Das A, Gurudu S, Sharma VK. Narrow Band Imaging (NBI) and White Light Endoscopy (WLE) have a comparable yield for detection of colon polyps in patients undergoing screening or surveillance colonoscopy: a meta-analysis. Gastrointest Endosc [Internet]. Nature Publishing Group; 2009;69:AB363. Available from: http://dx.doi.org/10.1038/ajg.2011.436.
51. Sabbagh LC, Reveiz L, Aponte D, et al. Narrow-band imaging does not improve detection of colorectal polyps when compared to conventional colonoscopy: a randomized controlled trial and meta-analysis of published studies. BMC Gastroenterol. 2011;11:100.
52. Nagomi A, Bjelakovic G, Petrovic B. Narrow band imaging versus conventional white light colonoscopy for the detection of colorectal polyps. Cochrane Database Syst Rev. 2012;CD008361.
53. Ramsoekh D, Haringsma J, Poley JW, van Putten P, van Dekken H, Steyerberg EW, et al. A back-to-back comparison of white light video endoscopy with autofluorescence endoscopy for adenoma detection in high-risk subjects. Gut. 2010;59:785–93.
54. Moriichi K, Fujiya M, Sato R, Watari J, Nomura Y, Nata T, et al. Back-to-back comparison of auto-fluorescence imaging (AFI) versus high resolution white light colonoscopy for adenoma detection. BMC Gastroenterol. [Internet]. BMC Gastroenterology. 2012;12:75. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3444423&tool=pmcentrez&rendertype=abstract.
55. van den Broek FJ, van Soest EJ, Naber AH, et al. Combining autofluorescence imaging and narrow-band imaging for the differentiation of adenomas from non-neoplastic colonic polyps among experienced and non-experienced endoscopists. Am J Gastroenterol. 2009;104:498–507.

# Appendix 4

# Example of Metaboanalyst data handling description


GC/LC-MS raw spectra
MS / NMR peak lists
MS / NMR spectra bins
Metabolite concentrations
• Peak detection
• Retention time correction
Peak alignment
Baseline filtering
Name/ID conversion
• Data integrity check
• Missing value imputation
DATA PROCESSING
Data normalization & transformation
DATA ANALYSIS
①Exploratory statistical analysis
②Functional enrichment analysis
③Metabolic pathway analysis
④Time-series and two factor analysis
⑤Sample size and power analysis
⑥ROC curve based biomarker analysis
⑦Gene-metabolite pathway analysis
⑧Utility functions
Downloads
• Processed data
• PDF report
• Result tables
• Images

# Metabolomic Data Analysis with MetaboAnalyst 3.0

User ID: guest844030029555876073

June 30, 2016

## 1 Data Processing and Normalization

### 1.1 Reading and Processing the Raw Data

MetaboAnalyst accepts a variety of data types generated in metabolomic studies, including compound concentration data, binned NMR/MS spectra data, NMR/MS peak list data, as well as MS spectra (NetCDF, mzXML, mzDATA). Users need to specify the data types when uploading their data in order for MetaboAnalyst to select the correct algorithm to process them. Table 1 summarizes the result of the data processing steps.

#### 1.1.1 Reading Concentration Data

The concentration data should be uploaded in comma separated values (.csv) format. Samples can be in rows or columns, with class labels immediately following the sample IDs.

Samples are in columns and features in rows. The uploaded file is in comma separated values (.csv) format. The uploaded data file contains 77 (samples) by 159 (compounds) data matrix.

#### 1.1.2 Data Integrity Check

Before data analysis, a data integrity check is performed to make sure that all the necessary information has been collected. The class labels must be present and contain only two classes. If samples are paired, the class label must be from -n/2 to -1 for one group, and 1 to n/2 for the other group (n is the sample number and must be an even number). Class labels with same absolute value are assumed to be pairs. Compound concentration or peak intensity values should all be non-negative numbers. By default, all missing values, zeros and negative values will be replaced by the half of the minimum positive value found within the data (see next section)

#### 1.1.3 Missing value imputations

Too many zeroes or missing values will cause difficulties for downstream analysis. MetaboAnalyst offers several different methods for this purpose. The default method replaces all the missing and zero values with a small values (the half of the minimum positive values in the original data) assuming to be the detection limit. The assumption of this approach is that most missing values are caused by low abundance metabolites (i.e.below the detection limit). In addition, since zero values may cause problem for data normalization (i.e. log), they are also replaced with this small value. User can also specify other methods, such as replace by mean/median, or use K-Nearest Neighbours, Probabilistic PCA (PPCA), Bayesian PCA (BPCA) method, Singular Value Decomposition (SVD) method to impute the missing values [1]. Please choose the one that is the most appropriate for your data.

[1] Stacklies W, Redestig H, Scholz M, Walther D, Selbig J. *pcaMethods: a bioconductor package, providing PCA methods for incomplete data.*, Bioinformatics 2007 23(9):1164-1167

Zero or missing variables were replaced with a small value: 0.5

#### 1.1.4 Data Filtering

The purpose of the data filtering is to identify and remove variables that are unlikely to be of use when modeling the data. No phenotype information are used in the filtering process, so the result can be used with any downstream analysis. This step can usually improves the results. Data filter is strongly recommended for datasets with large number of variables (> 250) datasets contain much noise (i.e.chemometrics data). Filtering can usually improve your results[2].

*For data with number of variables < 250, this step will reduce 5% of variables; For variable number between 250 and 500, 10% of variables will be removed; For variable number bwteen 500 and 1000, 25% of variables will be removed; And 40% of variabled will be removed for data with over 1000 varaibles.*

No data filtering was performed.

### 1.2 Data Normalization

The data is stored as a table with one sample per row and one variable (bin/peak/metabolite) per column. The normalization procedures implemented below are grouped into four categories. Sample specific normalization allows users to manually adjust concentrations based on biological inputs (i.e. volume, mass); row-wise normalization allows general-purpose adjustment for differences among samples; data transformation and scaling are two different approaches to make features more comparable. You can use one or combine both to achieve better results.

The normalization consists of the following options:

1. Sample specific normalization (i.e. normalize by dry weight, volume)
2. Row-wise procedures:
   - Normalization by the sum
   - Normalization by the sample median
   - Normalization by a reference sample (probabilistic quotient normalization)[3]
   - Normalization by a reference feature (i.e. creatinine, internal control)
3. Data transformation :
   - Generalized log transformation (glog 2)
   - Cube root transformation
4. Data scaling:
   - Unit scaling (mean-centered and divided by standard deviation of each variable)
   - Pareto scaling (mean-centered and divided by the square root of standard deviation of each variable)
   - Range scaling (mean-centered and divided by the value range of each variable)

Figure 1 shows the effects before and after normalization.

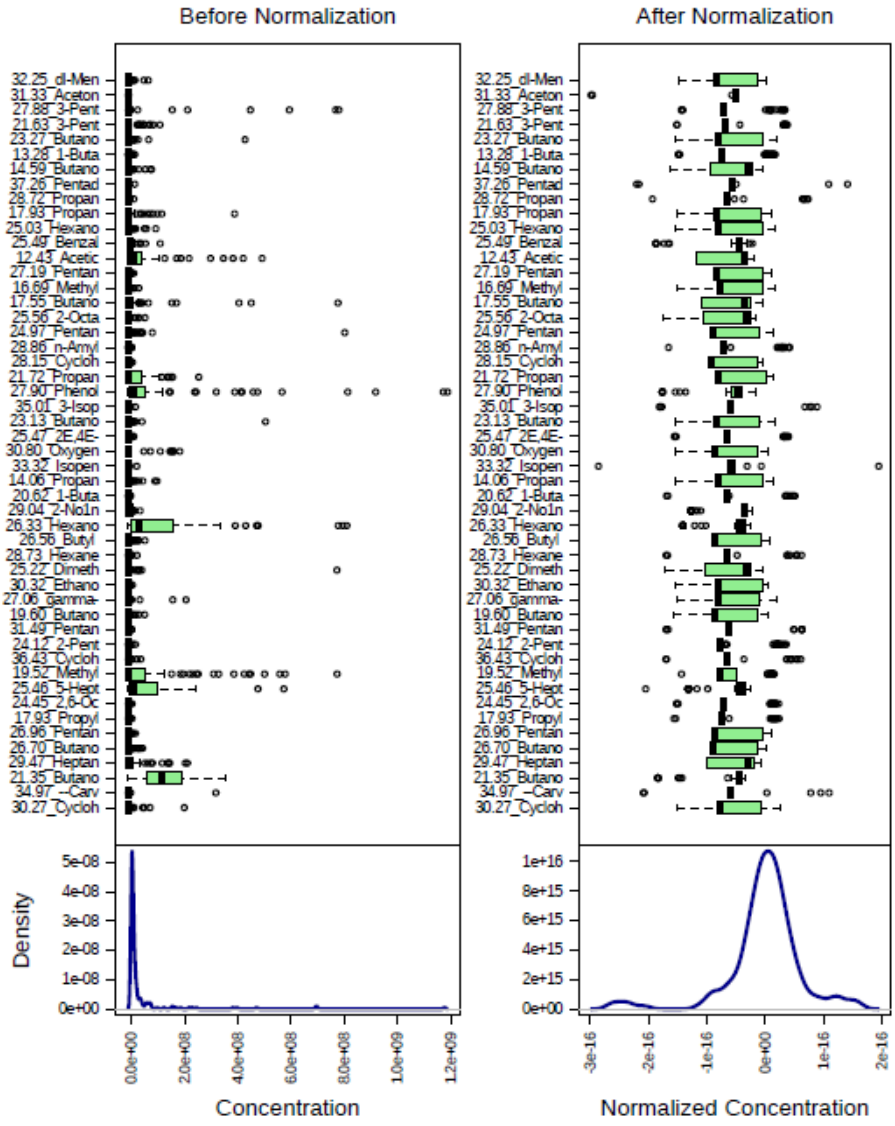



Figure 1: Box plots and kernel density plots before and after normalization. The boxplots show at most 50 features due to space limit. The density plots are based on all samples. Selected methods : Row-wise normalization: Normalization to sample median; Data transformation: Log Normalization; Data scaling: Autoscaling.

## 2 Statistical and Machine Learning Data Analysis

MetaboAnalyst offers a variety of methods commonly used in metabolomic data analyses. They include:

1. Univariate analysis methods:
   - Fold Change Analysis
   - T-tests
   - Volcano Plot
   - One-way ANOVA and post-hoc analysis
   - Correlation analysis
2. Multivariate analysis methods:
   - Principal Component Analysis (PCA)
   - Partial Least Squares - Discriminant Analysis (PLS-DA)
3. Robust Feature Selection Methods in microarray studies
   - Significance Analysis of Microarray (SAM)
   - Empirical Bayesian Analysis of Microarray (EBAM)
4. Clustering Analysis
   - Hierarchical Clustering
     - Dendrogram
     - Heatmap
   - Partitional Clustering
     - K-means Clustering
     - Self-Organizing Map (SOM)
5. Supervised Classification and Feature Selection methods
   - Random Forest
   - Support Vector Machine (SVM)

```
Please note:  some advanced methods are available only for two-group sample analyais.
```

## 3 Data Annotation

Please be advised that MetaboAnalyst also supports metabolomic data annotation. For NMR, MS, or GC-MS peak list data, users can perform peak identification by searching the corresponding libraries. For compound concentration data, users can perform metabolite set enrichment analysis and metabolic pathway analysis.

---

The report was generated on Thu Jun 30 09:38:59 2016 with R version 3.2.2 (2015-08-14). Thank you for using MetaboAnalyst! For suggestions and feedback please contact Jeff Xia (*jeff.xia@mcgill.ca*).