WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/131065/

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

Advances in the detection of colorectal cancer

By Ella Mozdiak

MBChB, MRCP (UK)

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Medicine

School of Engineering and Warwick Medical School,

University of Warwick

August 2018

Table of contents

Abbreviations ……………………………………………………………...8

List of figures …………………………………………………………. 13

List of tables ..……………………………………………………………. 16

Acknowledgements. ……………………………………………………… 18

Declaration …………………………………………………………………. 20

List of publications ……………………………………………………… 21

Abstract …………………………………………………………………….. 22

Chapter 1. Introduction ……………………………………………… 23

1.1 Clinical scope …………………………………………………………. 24

1.2 Research Aims ……………………………………………………….. 27

1.3 Thesis structure ……………………………………………………… 27

1.4 Conclusion ……………………………………………………………29

Chapter 2. Literature review ………………………………………… 30

2.1 The anatomy of the colon ……………………………………………31

2.2 Colonic polyps ………………………………………………………...31

2.2.1 Colonic polyposis syndromes and Lynch syndrome ……………….. 35

2.3 Colorectal cancer ………………………………………………………37

2.3.1 Pathology ……..…………………………………………………37

2.3.2 Staging …………….……………………………………………38

2.3.3 Aetiology ……………………………………………………… 39

2.3.3.1 Diet ……………………………………………..……………39

2.3.3.2 Lifestyle factors ………………………………………………....40

2.3.3.3 The microbiome ……………………………………………...40

2.3.3.4 Inflammatory bowel disease ……………………......................41

2.3.4 Chemoprevention …………………………………………………. 42

2.3.5 Treatment ………………………………………………………. 43

2.3.5.1 Surgery ……………………………………………………...43

2.3.5.2 Chemotherapy …………………………………………………... 44

2.3.5.3 Radiotherapy ……………………………………………………. 44

2.3.5.4 Other therapy ……………………………………………………......44
2.4 Diagnosis of colorectal cancer ………………………………………...45
2.4.1 Tumour and bio markers ………………………………………….. 46
2.5 Diagnostic pathways in colorectal cancer ……………………………49
2.5.1 Bowel cancer screening …………………………………………….. 49
2.5.1.1 The screening process ………………………………………… 50
2.5.1.2 The drawbacks of using the guaiac FOBT ……………………...52
2.5.1.3 Faecal immunochemical testing (FIT) ………………………… 53
2.5.1.4 Endoscopic population screening ……………………………… 54
2.5.2 The two-week wait (TWW) referral system ……………………… 55
2.6 Volatile organic compounds (VOCs) …………………………………57
2.6.1 The origin of VOCs ……………………………...……. ……..58
2.6.2 Instruments used in VOC analysis …………………………………59
2.6.3 VOC detection in non-gastrointestinal disease ……………………63
2.6.4 VOCs in gastrointestinal disease …………………………………. 68
2.6.5 VOCs in the detection of colorectal cancer ………………………. 70
2.6.5.1 VOC profile and cancer stage …………………………………75
2.7 The proteome …………………………………………………………76
2.7.1 Proteomics and cancer detection ………………………………….. 77
2.7.2 Instruments used in protein analysis ………………………………80
2.8 Conclusion ……………………………………………………………..83
Chapter 3:
The effect of the two-week wait referral system on the detection of colorectal cancer: a systematic review and meta-analysis ……….…..84
3.1 Introduction …………………………………………………………… 85
3.2 Objectives ………………………………………………………………86
3.2.1 Primary outcomes …………………………………………………... 86
3.2.2 Secondary outcomes …………………………………..……….86
3.3 Method …………………………………………………………………87
3.3.1 Study eligibility …………………………………………………. 87
3.3.2 Study inclusion criteria ……………………………………………87
3.3.3 Study exclusion criteria ……………………………………………...88

3.3.4 Search strategy 89
3.3.5 Study selection 89
3.3.6 Data extraction 90
3.3.7 Quality assessment 90
3.3.8 Data analysis and synthesis 91
3.3.9 Deviations from the protocol 92
3.4 Results 93
3.4.1 Search results 93
3.4.2 Studies and their characteristics 95
3.4.3 Population characteristics 95
3.4.4 Studies including a subset population 109
3.4.4.1 Age-restricted cohort 109
3.4.4.2 Symptom-restricted and image-restricted cohorts 110
3.4.5 Quality of included studies 110
3.4.6 Synthesis of outcome data 110
3.4.7 Cancer conversion rate 111
3.4.8 Cancer stage data from studies with no comparator group............113
3.4.9 Cancer stage from studies with a comparator
(non-TWW) group 116
3.4.10 Other information on cancer stage 117
3.4.11 Other diagnosis 117
3.4.12 Polyp detection rate 118
3.4.13 Non-CRC malignancy and other GI diagnoses 118
3.4.14 Adherence 121
3.4.15 Survival 122
3.4.16 Assessment of heterogeneity and publication bias 123
3.5 Discussion 124
3.5.1 Cancer conversion rate and stage at diagnosis 124
3.5.1.1 Major findings 124
3.5.1.2 Strengths and limitations of the evidence 125
3.5.1.3 Consistency of the findings 125
3.5.2 Other non-CRC diagnoses made and adherence to

referral targets ……………………………………………………….…..126
3.5.2.1 Major findings ………………………………………………126
3.5.2.2 Strengths and limitations of the evidence …………. ……………127
3.5.2.3 Consistency of findings ………………………………………127
3.5.3 Survival ………………………………………………………….. 127
3.5.3.1 Strengths and limitations of the evidence …………………………128
3.5.4 Impact of the results …………………………………………….. 128
3.5.5 Strengths and limitations of the review …………………………….. 132
3.5.6 Future research recommendations …………………………………. 132
3.6 Conclusion ……………………………………………………………..134
Chapter 4: Methodological development of volatile organic compound analysis in the detection of colorectal cancer…………………………… 135
4.1 Introduction………………………………………………………… 136
4.2 Analytical techniques………………………………………………….. 137
4.2.1 FAIMS………………………………………………….……137
4.2.2 FAIMS matrix ………………………………………………….143
4.2.3 Statistical analysis …………………………………………………...145
4.3 GC-IMS analysis ……………………………………………………..148
4.3.1 Manual sampling method ……………………………………………151
4.4 Experimental work …………………………………………………… 151
4.4.1 Experiment one: urine validation …………………………………... 152
4.4.2 Experiment two: CRC and controls …………………………………155
4.4.3 Experiment three: sample storage ………………………………….. 156
4.4.4 Experiment four: urine degradation …………………………………158
4.5 Discussion ……………………………………………………………... 167
4.6 Conclusion and further work ………………………………………….. 169
Chapter 5: The application of urinary volatile organic compound analysis in the detection of colorectal cancer……………….............. 170
5.1 Introduction ………………………………………………………… 171
5.2 Hypothesis ……………………………………………………………172
5.3 Aims …………………………………………………………………… 172
5.4 Research questions …………………………………………………….. 172

5.5 Ethical approval ……………………………………………………….172
5.6 Machine Learning ………………………………………………………173
5.7 Experimental work ……………………………………………………...173
5.7.1 Experiment one: The application of urinary VOC analysis in the detection and differentiation of CRC by tumour stage in the symptomatic population: A retrospective case-control study……....173
5.7.2 Experiment two: A prospective cohort study in the BCSP population, comparing urinary VOC profiles with diagnostic outcomes: (phase one) ………………………………………………177
5.7.3 Experiment three: A prospective cohort study in the bowel cancer screening population comparing urinary VOC profiles with diagnostic outcomes (phase two) ……..........................185
5.8 Discussion …………………………………………………………......196
5.9 Conclusion ……………………………………………………………..202
Chapter 6: Investigating the urinary proteome in the detection of colorectal cancer ……………………………………………………203
6.1 Introduction …………………………………………………………….. 204
6.2 Aims ……………………………………………………………………. 205
6.3 Research questions …………………………………………………….. ..205
6.4 Ethical approval …………………………………………………………205
6.5 Experimental work ……………………………………………………… 206
6.5.1 Identification of urinary peptide markers for colorectal cancer by capillary electrophoresis coupled to mass spectrometry ….206
6.5.2 Experiment two: Establishing the role of urinary-derived proteins in the pathogenesis of CRC …………………………… 218
6.6 Discussion ………………………………………………………………. 226
6.6.1 Returning to the research questions initially posed …………………. 230
6.7 Conclusion ………………………………………………………………231
Chapter 7: Discussion ……………………………………………………232
7.1 Introduction …………………………………………………………….. 233
7.2 The detection of CRC using the TWW referral pathway …………….... 233
7.3 Biomarker development ………………………………………………… 235

7.4 VOCs in the detection of CRC ……………………………………………...238
7.5 Proteomics in the detection of CRC …………………………………….. 239
7.6 The future of biomarker research in CRC ………………………………..241
7.7 Recommendations and further research ………………………………….243
7.8 Limitations ………………………………………………………………….243
7.9 Final remarks ……………………………………………………………….244
Bibliography ………………………………………………………………….. 245
Appendices ……………………………………………………………………. 275

Abbreviations

^{0}C degrees centigrade

Ml millilitres

m/s metres per second

µg microgram

µl microliter

µmol micromole

18Qloh 18q Loss of heterozygosity

ADR Adenoma detection rate

AJCC American Joint Committee on cancer

APC Adenomatous polyposis coli

ASA Aminosalicylic acid

ATLAS At-line sampling module

ATPase Adenosine triphosphatase

AUC Area Under the Curve

BCSP Bowel cancer screening programme

BMI Body mass index

BSG British Society of Gastroenterology

CA 19-9 Cancer antigen 19-9

CD Crohn’s disease

CEA Carcinoembryonic antigen

CEBIS Clinical Evidence Based Information Service

CE-MS Capillary electrophoresis mass spectrometry

CHRPE Congenital hypertrophy of the retinal pigment epithelium

CI Confidence interval

COPD Chronic obstructive airways disease

COX Cyclo-oxygenase

CRF Case report form

CRP C-reactive protein

CTC Computed tomograpic colonography

CV Compensation voltage

DAB diaminobenzidine

DC Direct current

DCC Deleted in colon cancer (gene)

DFMO difluoromethylornithine

DMS Differential mobility spectrometry

DNA Deoxyribonucleic acid

EGFR epidermal growth factor receptor inhibitor

ELISA Enzyme-linked immunosorbent assay

ESI Electrospray ionisation

FAIMS Field asymmetric ion mobility spectrometry

FAMISHED Food and fermentation using metagenomics in health and disease

FAP Familial Adenomatous Polyposis

FCP Faecal Calprotectin

FDA Fisher discriminant analysis

FEV Forced expiratory volume

FOBT Faecal occult blood test

FODMAP fermentable oligosaccharides, disaccharides, monosaccharides and polyols

FIT Faecal Immunochemical test

FTMS Fourier transform mass spectrometer

FU Fluorouracil

GC-MS Gas chromatrography- mass spectrometry

GC-IMS Gas chromatography ion mobility spectrometer

GI Gastrointestinal

GORD Gastro-oesophageal reflux disease

GP General Practitioner

HMG-CoA 3-hydroxy-3-methylglutaryl-coenzyme A

HNPCC Hereditary non-polyposis colorectal cancer

HPLC High performance liquid chromatography

IBD Inflammatory bowel disease

IBS Irritable bowel syndrome

IHS Immunohistochemistry

IMA Inferior mesenteric artery

IPI International Protein Index

JAG Joint Advisory Group

KRAS Kirsten rat sarcoma virus

LC Liquid chromatography

MDT Multi-disciplinary team

MEP1A Meprin alpha

MFC Mass flow control

MMP Matrix metalloproteases

MMR Mismatch repair

MSI Microsatellite instability

MUTYH mutY homolog

NCIN National cancer intelligence network

NHS National Health Service

NICE National Institute for Health and care excellence

NIST National institute of standards and technology

NPV Negative predictive value

NSAIDS Non-Steroidal Anti inflammatory Drugs

NSC National screening committee

PAGE polyacrylamide gel electrophoresis

PCA Principle component analysis

PCP Planar cell polarity

PCR Polymerase chain reaction

PIP Product ion peak

PPV Positive predictive value

PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PROSPERO International prospective register of systematic reviews

PSA Prostate specific antigen

RF Radiofrequency

RCC Renal cell carcinoma

RIP Reactive ion peak

ROC Receiver operating characteristic

RNA Ribonucleic acid

RR – Relative risk

SELDI surface-enhanced laser desorption and ionisation

SIFT-MS Selected Ion Flow Tube Mass Spectrometry

SLT Sparse logistic regression

SMOTE Synthetic Minority Over Sampling Technique

SMA Superior mesenteric artery

SPME Solid phase micro-extraction

SPS Serrated polyposis syndrome

SVM Support vector machine

Taa Tumour associated antigen

TBS Tris-buffered saline

T2DM Type 2 Diabetes Mellitus

Tp53 Tumour suppressor p53

TRIP Turning research into practice

TWW Two-week wait

UC Ulcerative colitis

UHCW University Hospitals Coventry and Warwickshire

UK United Kingdom

UTI Urinary tract infection

Wnt Wingless type

VEGF Vascular endothelial growth factor

VOC Volatile Organic Compounds

WHO World Health Organisation

List of figures

Chapter 1

1.1 CRC incidence and mortality

1.2 Proportion of CRCs Diagnosed at Each Stage

1.3 Age-Standardised CRC Five-Year Relative Survival

Chapter 2

2.1 Anatomy of the colon with labelled constituent parts

2.2 A picture of a pedunculated adenomatous rectal polyp

2.3 The adenoma-carcinoma sequence

2.4 The BSG CRC screening and surveillance flowsheet

2.5 A picture of a colonic adenocarcinoma

2.6 A picture of the guaiac faecal occult blood test

2.7 A picture of the faecal immunochemical test kit

2.8 A flow diagram of the TWW pathway

2.9 Schematic of the comparison between the human and electronic nose

2.10 Schematic of a GC-MS instrument

Chapter 3

3.1 PRISMA flow diagram of study selection process

3.2 A forest plot of cancer conversion rate

3.3a and b Forest plot of papers reporting proportions presenting at Dukes stages A-D

3.4 Forest plot of adenomas and undefined polyp detection rate

3.5 Forest plot of non-CRC diagnosis rates

Chapter 4

4.1 Schematic of the FAIMS analysis process

4.2 Picture of the ATLAS sampling module

4.3 Picture of the Lonestar FAIMS instrument together with auto sampler

4.4 Lonestar colourplot showing a dispersion field matrix

4.5 Lonestar dispersion field matrix demonstrated by a waterfall plot

4.6 Schematic of the separation process and ion detection using GC-IMS

4.7 The two-stage separation of a biological sample using GC-IMS

4.8 3-Dimensional representation of GC data output with corresponding IMS chromatogram
4.9 An abnormal Lonestar plume
4.10 Scatter plot showing change in total number of urinary VOCs over time
4.11 Scatter plot showing chemical diversity of urinary VOCs over time
4.12 Chart of variation of FAIMS matrices as an average of all samples with increasing time
4.13 Variation in FAIMS matrices as seen by plume structure
4.14 Chart of trend in total ion count from mean of FAIMS matrices
4.15a Positive FAIMS matrices: variation in ion count from baseline as a function of matrix number
4.15b Negative FAIMS matrices: variation in ion count from baseline as a function of matrix number
4.16a Positive FAIMS matrices: variation from baseline as a function of matrix number
4.16b Negative FAIMS matrices: variation from baseline as a function of matrix number
4.17 Positive FAIMS matrices: difference from baseline with matrix number
Chapter 5
5.1 ROC curve for CRC vs control in the symptomatic population
5.2 ROC curve for CRC vs no disease in the BCSP cohort
5.3 ROC curve for classification of CRC vs normal in BCSP cohort
5.4 ROC curve of classification of CRC vs normal using GC-IMS
5.5a GC-IMS raw output data before feature selection
5.5b GC-IMS raw output data after feature selection
Chapter 6
6.1 Simple schematic of the CE-MS process
6.2 Identification chain from urinary peptides to proteins
6.3 Contour plots of the raw CE-MS data for CRC, adenomas and control
6.4 ROC curve for differentiating CRC from non-CRC using a 26 urinary peptide marker model

Figure 6.5a Digital image of IHC staining of MEP1A in the normal colonic tissue epithelium of case number 6
Figure 6.5b Digital image of IHC staining of MEP1A in the CRC tissue of case number 6
Figure 6.5c Digital image of IHC staining of MEP1A in the advancing margin of CRC tissue in case number 8
Chapter 7
7.1 Schematic of the factors influencing potential biomarkers
7.2 Five phases of biomarker development

List of tables

Chapter 1

1.1. Five year relative CRC survival by stage

Chapter 2

2.1 Summary table of publications reporting VOC biomarker use in detection of CRC

2.3 Approved purely protein-based biomarkers in cancer

2.4 Analytical techniques currently available for proteomics

Chapter 3

3.1 Study and population characteristics

3.2 Summary of all papers reporting only TWW subset data

3.3 Staging data from comparator (non-TWW) groups

3.4 Summary of all data reported on adherence to operational standards for TWW referrals

3.5 Studies that reported survival rates in the TWW vs non-TWW cohorts

Chapter 4

4.1 Study population characteristics of the urine validation cohort

4.2 Ion concentration data per FAIMS matrix

Chapter 5

5.1 Study population characteristics for CRC and IBS groups

5.2 Cancer distribution by site

5.3 Stage of cancer according to Dukes criteria

5.4 Study population characteristics and medications for BCSP patients

5.5 Diagnostic outcomes in the BCSP cohort

5.6 Classification of BCSP cohort by outcome

5.7 Diagnostic outcomes for BCSP cohort (phase two)

5.8 Classification of BCSP cohort (phase two) by outcome

5.9 Classification of BCSP cohort: CRC vs adenoma

5.10 Classification of BCSP cohort using GC-IMS

Chapter 6

6.1 Study population characteristics

6.2 Proteins from which the CRC peptide marker candidates in urine are derived

6.3 CE-MS and sequence characteristics of the 26 CRC peptide markers candidates that were detected by CE-MS in plasma and urine
6.4 Classification of CRC from control using a 26-peptide urinary marker tool
6.5 Nine antigens (proteases) which have a link with the raw urinary peptide data
6.6 Allred scoring system for stain intensity
6.7 Tumour site and stage
6.8 IHC staining intensity score for MEP1A in CRC compared with normal tissue

Acknowledgements

Firstly I would like to express my gratitude to my clinical supervisor, Professor Ramesh Arasaradnam for having faith in me and showing unwavering encouragement and enthusiasm to help develop my research skills and knowledge.

Also I am grateful to my academic supervisors, first Professor Mark Pallen and subsequently Professor Judith Klein.

This work has relied on the dedication of the gastroenterology research nurses at UHCW, namely Nicola O'Connell, Subiatu Wurie and Leanne Burns who have tirelessly recruited patients. Also to the bowel cancer screening nurses, Paula Voisey and Carol Wheatley.

My thanks also go to the whole team in the Biosensors lab at The University of Warwick. To Professor James Covington for having patience with a medic in his laboratory and for providing support with all practical aspects of this work. Also to Emma Daulton, Siavash Esfahani and Alfian Wickasono for their invaluable help with sample and data analysis. I am very grateful to Professor Rich Savage and Matt Neal for their help with data analysis.

I am also indebted to Yonas Weldeselassie and Alexander Tservadze for their invaluable expert assistance with the systematic review.

The tissue bank team at UHCW have been hugely supportive, these include Parmjit Dahaley, Adrian Fisk and Andrew White. Special thanks go to Sean James for his patience in teaching me immunohistochemistry. Thank you to Dr Kishore Gopalakrishnan for providing his expertise in pathology.

To Mosaiques Diagnostics for performing CE-MS and in particular, to Jochen Metzger for his assistance in evaluating the data.

To the Medical and Life Sciences Research Fund and to The Midlands Gastroenterological Society for their generous research grants.

I wholeheartedly thank all the patients who kindly donated specimens in the aid of this research.

Finally, to my family: To my mother and father for their encouragement and endless proof-reading. To my husband for his support, reassurance and constant belief in my ability. Finally to the two biggest inspirations in my life: my boys, Joseph and Sebastian, who put everything else into perspective.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Medicine. I, Ella Mozdiak, declare that it has been composed by myself and has not been submitted in any previous application for any degree and all the research has been undertaken in accordance with University safety policy and guidelines on ethical practice.

The presented work, including data generation and analysis, was carried out by the author except in the cases outlined below:

- Statistical analysis of the FAIMS data was performed by Matt Thomas and Matt Neal (Chapter Four)
- Statistical analysis of the FAIMS and GC-IMS data was performed by Alfian Wickasono (Chapter Five)
- Urinary peptide identification using CE-MS was performed commercially by Mosaiques Diagnostics (Chapter Six)

List of publications

1. Mozdiak E, Tsersvadze A, McFarlane M, Widlak M, Tabuso M, Dunlop A, Arasaradnam R. The effect of the 2-week wait referral system on the detection of and mortality from colorectal cancer: protocol of a systematic review and meta-analysis. Syst Rev. 2016;5(1):182

2. Esfahani S, Sagar NM, Kyrou I, Mozdiak E, O'Connell N, Nwokolo C, Bardhan KD, Arasaradnam RP, Covington JA. Variation in Gas and Volatile Compound Emissions from Human Urine as it Ages, Measured by an Electronic Nose. Biosensors (Basel). 2016; 6(1).

Abstract

Colorectal cancer (CRC) is a significant cause of mortality. In the UK, 1 in 14 men and 1 in 19 women will develop CRC in their lifetime. Overall, five-year survival is 60% and this compares poorly to many countries in Western Europe. Survival is directly linked to stage at presentation. The majority present at Dukes stage B or above. Better methods of detecting CRC at an earlier stage are needed to improve survival figures.

Published evidence on the two-week wait pathway (TWW) for detecting CRC was evaluated using a systematic review. This showed low CRC detection and no difference in stage at presentation in the TWW group compared with the non-TWW pathways of referral. These results indicate the emphasis on TWW targets is not justified and should prompt consideration of better methods of risk stratification for those with suspected CRC.

Biomarkers have the potential to improve the positive predictive value of referral criteria in the symptomatic population and also increase the sensitivity of the faecal occult blood test when used as an adjunct in the bowel screening population. Two potential urinary biomarker groups in the detection of CRC were explored: one based on volatile organic compounds (VOCs) and one on peptide markers. Analysis demonstrated that CRC could be correctly classified from control using a VOC or peptide biomarker-based test with a high degree of accuracy in pilot work. Further protein analysis using immunohistochemistry explored the role of Meprin alpha in the pathology of CRC with possible implications in aggressive disease.

There is a need for better ways of detecting CRC and biomarkers hold huge potential. But there must be a culture change from the current fragmented approach to elucidating diagnostics in CRC. Collaborations between research groups and data-sharing is essential to validate pilot studies and share knowledge on methodology and results.

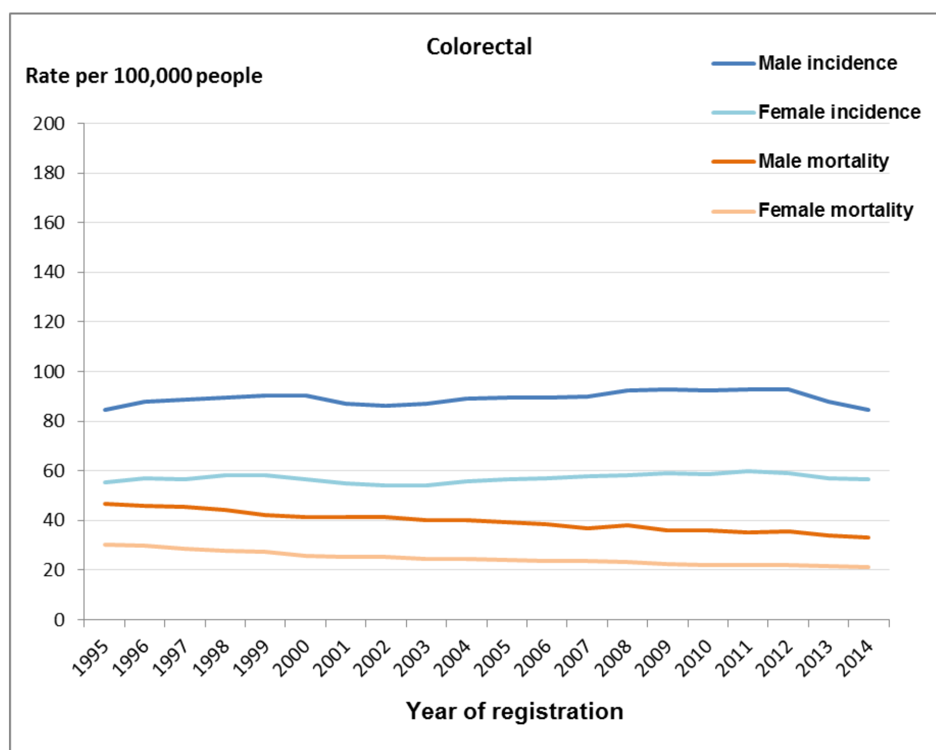
CHAPTER 1:

Introduction

1.1 Clinical Scope

Colorectal cancer (CRC) is a leading cause of cancer-related deaths in the Western world with highest rates in North America, Europe and Australia (1, 2). It represents the fourth most common cancer in England and Wales and is the second leading cause of cancer-related deaths(1). There were 41,265 new cases in 2014, and it makes up over 10% of all cancer diagnoses in the England and Wales population. Like many cancers, CRC predominantly affects older age groups, especially people aged 60 years or older, with peak rates of 85-89 years (3). The incidence has remained relatively consistent over the past 20 years, although there has been a downward trend in mortality for both males and females (figure 1.1) (4).

Figure 1.1. CRC incidence and mortality. *Directly age-standardised rates per 100,000 population of newly diagnosed cases of CRC and deaths from cancer (4).*



Currently, the overall five-year survival is approximately 50% (5, 6). Survival is high for early stage disease but declines sharply with advancing stage at diagnosis as shown in table 1.1 (3).

Table 1.1. Five year relative CRC survival by stage *for adults in England, 2002-2006 (3).*

Stage	Men	Women
Stage I	94.6	100.2
Stage II	83.5	85.9
Stage III	62.6	62.7
Stage IV	6.9	8.1
All Stages	58.2	61.1
Stage Not Known	18.7	15.1

Despite rising awareness of CRC, the majority of patients present at Dukes stage B/stage ii or above, this is represented graphically in figure 1.2. The United Kingdom (UK) survival figures fall short of many of our Western European counterparts (5, 7, 8). Wales has the poorest survival in the UK with figures similar to Estonia and Croatia rather than Germany and Switzerland (figure 1.3). It has been estimated that 10,000 deaths could be avoided each year if the cancer mortality figures were similar to the lowest rates in Europe (6).

Figure 1.2. Proportion of CRCs Diagnosed at Each Stage*, All Ages, England 2014 (3)*

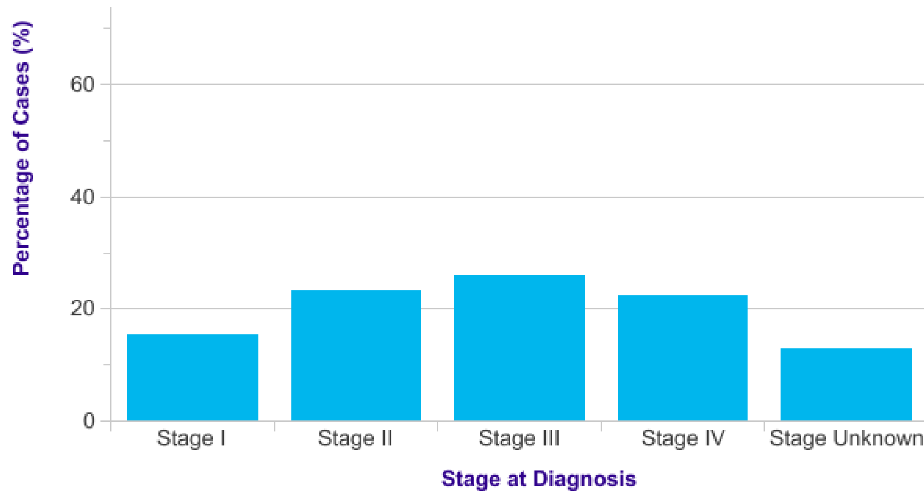
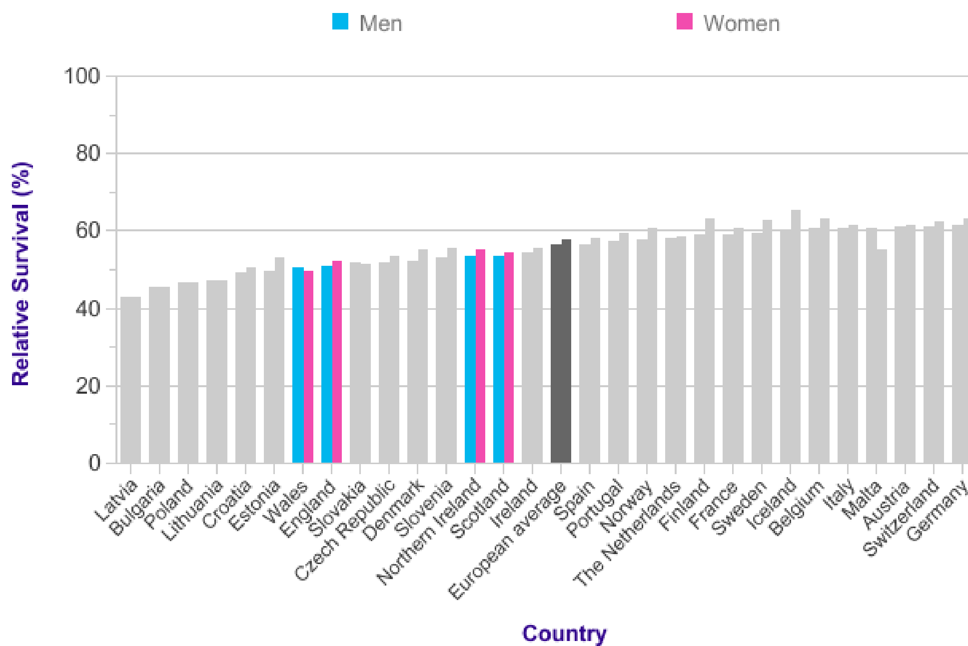


Figure 1.3. Age-Standardised CRC Five-Year Relative Survival, Adults in European Countries, 2000-2007 (9)



Although it is a complex issue, the UK's poor performance in Europe is thought, in part, to be related to a disproportionately high number of patients presenting at an advanced stage or as an emergency and possibly due to delays in treatment (10). Approximately 25 % of patients with CRC present as an emergency: this proportion

rises in the elderly and is higher than in many other common cancers (3). This has a direct effect on one-year survival rates: 48% in those presenting as emergencies compared with 73% through other routes of presentation (5). The early recognition of symptoms and prompt diagnosis are essential in reducing mortality (11).

The financial burden associated with CRC is significant. It is estimated that total CRC cost amounts to £1.6 billion per year in the UK. This includes the economic, healthcare and the unpaid care costs provided by family and friends (12).

In view of the impact of CRC and the need to address survival figures, this thesis re-evaluates the referral pathways for CRC, using a systematic review and meta-analysis. The emerging advances in the diagnosis of CRC, are investigated through a series of practical experiments. This work is intended to demonstrate progression in the understanding of the diagnostic options in this disease and set the foundations for further experimental work in bio-marker diagnostics.

1.2 Research Aims

1. To comprehensively evaluate the evidence reported on CRC detection using the two-week wait (TWW) referral pathway.
2. To perform a detailed analysis of the methods used for urinary volatile organic compound (VOC) detection in CRC.
3. To evaluate the utility of urinary VOC detection-based technology as a diagnostic tool in CRC applied to both the symptomatic and asymptomatic population.
4. To evaluate the utility of urinary peptide markers as a diagnostic tool in CRC and evaluate the role of target proteins in the pathogenesis of the disease.

1.3 Thesis structure

Chapter Two is a literature review which explains the pathophysiology, diagnosis and treatment of CRC as this disease is central to the thesis. The current UK routes

to diagnosis of the disease are highlighted to provide the background for the work carried out in Chapter Three. The later part of the literature review focusses on the theory and research into the two areas of CRC biomarker development that form the experimental work in this thesis, namely VOCs and peptide markers.

In Chapter Three the results of a systematic review and meta-analysis are presented. This chapter investigates, for the first time, the key outcome measures in the fast-track referral of patients with possible CRC. The impact of these results is explored and recommendations for improvement and development of the pathway are detailed

Chapter Four is the first of two chapters investigating the volatile content of urine headspace for VOCs. This chapter provides an in-depth analysis of the methodology used in this relatively new technology with a series of experiments that contributed improvements in urine sampling and analysis. It provides a hands-on guide for other researchers also utilising this technology.

The experimental work on VOCs is continued in Chapter Five with focus on CRC cases. Several experiments involving symptomatic and asymptomatic patients with the disease are presented. These are the first results to present data on the bowel cancer screening population.

Another relatively new and exciting technology in CRC diagnosis using urine is evaluated in Chapter Six. Firstly, capillary electrophoresis coupled to mass spectrometry is used to detect peptide markers unique to CRC and a peptide-based diagnostic test is proposed. Following sequencing of these peptides, the role of the corresponding proteins are considered in the pathology CRC.

The results of the tested technologies and diagnostic pathways and the impact these have on CRC diagnosis are debated in the concluding chapter and suggestions are made for ongoing research.

1.4 Conclusion

The impact of CRC on morbidity and mortality is significant. Research into more effective ways of diagnosing patients, particularly at an earlier stage, is vital to improving the survival figures. The following four results chapters aim to approach different aspects of CRC diagnosis. From pathway development, for streamlining symptomatic patients, to advances in laboratory based biomarkers using the newest emerging techniques that could have a direct benefit to patients.

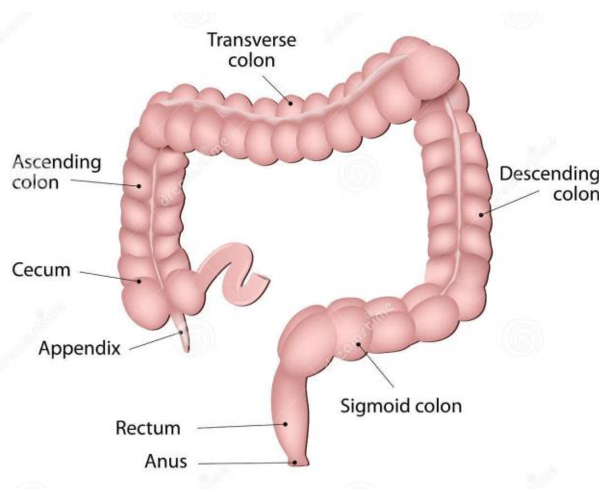
CHAPTER 2:

Literature review

2.1 The anatomy of the colon

The colon, or large intestine, is situated in the lower abdomen. It is the final part of the digestive tract and forms the connection between the small intestine and the anal canal. It is comprised of the following sections; Caecum, Ascending colon, Transverse colon, Descending colon, Sigmoid colon and Rectum (figure 2.1) and measures around 150cm in length (13). Blood supply arises from the Superior Mesenteric Artery and Inferior Mesenteric Artery. The colonic wall is composed of four layers: the mucosa, submucosa, muscularis and adventitia. The surface epithelium is simple columnar cells (14). The colon functions mainly to absorb water and acts as a storage compartment for faeces prior to evacuation.

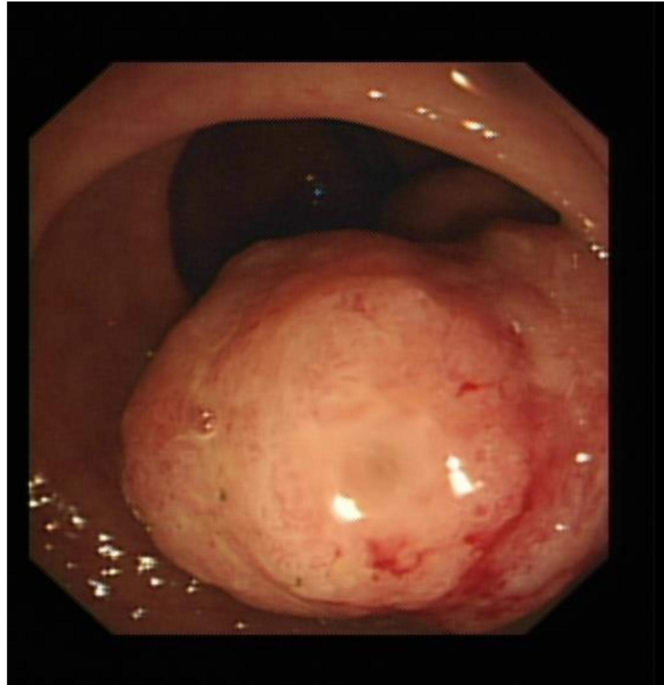
***Figure 2.1 Anatomy of the colon with labelled constituent parts** (15).*



2.2 Colonic polyps

Polyps are a mass of tissue that protrudes into the colonic lumen. The characterisation of polyps relates to their size, whether they have a stalk (pedunculated) or are flat (serrated) and their overall appearance (figure 2.2). In general terms polyps can be divided into neoplastic (adenomas and carcinomas) and non-neoplastic (hyperplastic, juvenile,Peutz-Jeghers and Inflammatory) (13). Most polyps are asymptomatic, although due to their protuberance, they can ulcerate, causing pain or bleeding.

Figure 2.2 A picture of a pedunculated adenomatous rectal polyp *(patient permission granted).*

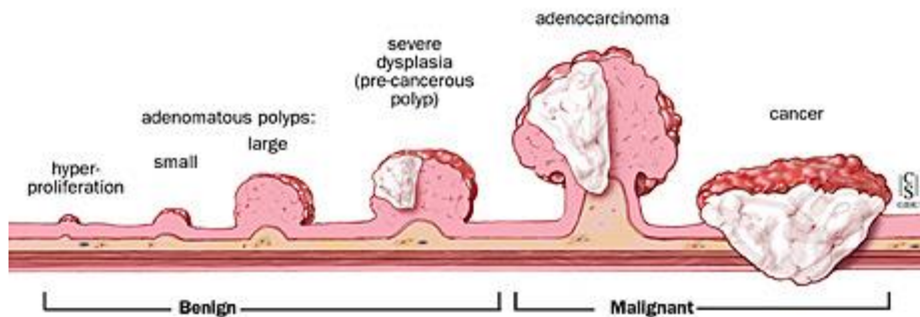


Adenomas are the most common neoplastic polyp and are by definition dysplastic (16). Dysplasia is defined as an alteration in size, shape, and organization of an adult cell. Most CRC arises from adenomas, but most adenomas have a benign course, with only five percent progressing to cancer (13, 16). Adenomas are more common in men and with increasing age, affecting up to 50% of those over 70 years of age (16, 17). They are commonly categorised by size into three groups: less than 1cm, 1-2cm and greater than 2cm. Most are less than 1cm in size, increasing size is associated with a higher risk of progression to cancer. Other factors determining the potential to progress to cancer are histological type and degree of dysplasia (18). It is thought that progression from adenomatous polyp to carcinoma in susceptible individuals takes at least five years (18).

Adenomatous polyps are thought to arise due to a disruption in the usual process of cell proliferation and cell death. It is uncontested that most carcinoma arises from

adenomas (19). The adenoma-carcinoma sequence describes the progression from normal epithelium to dysplastic and then subsequent carcinoma (figure 3).

Figure 2.3. The adenoma-carcinoma sequence *(adapted from Fearon and Vogelstein) (20)*



Developed by Fearon and Vogelstein (20), there is a robust body of clinical, epidemiological, pathological and molecular genetic evidence to support the adenoma-carcinoma sequence:

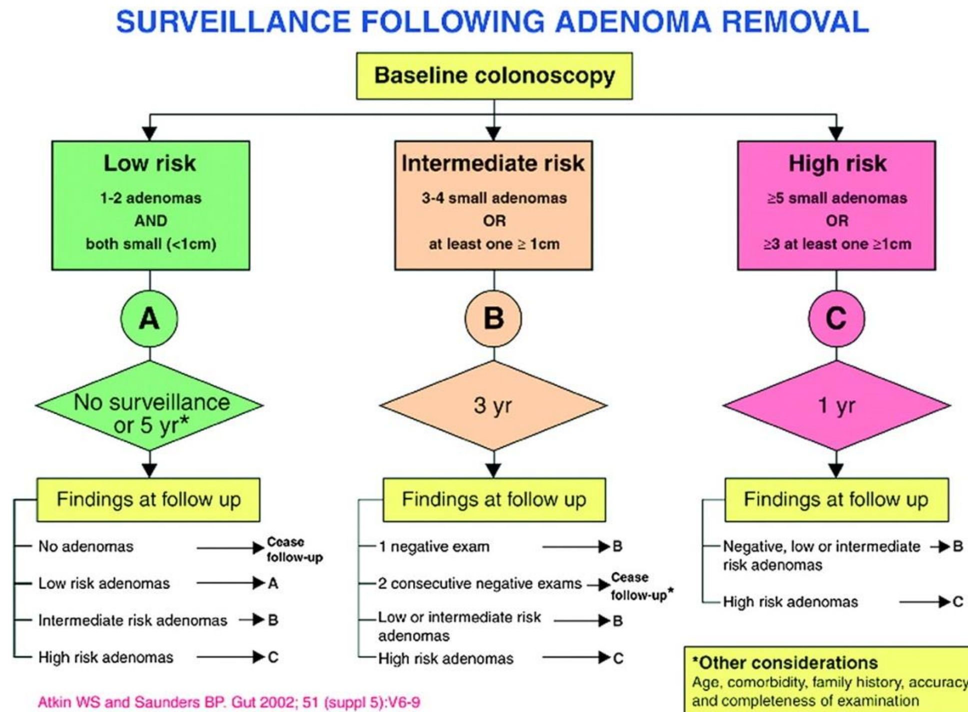
1. The prevalence of adenomas in a population parallels the incidence of CRC (19).
2. The anatomical distribution of adenomas and cancers appears to be similar, with the majority found distal to the splenic flexure, in the left colon (21).
3. Patients with polyposis syndromes such as Familial Adenomatous Polyposis (discussed later) will develop CRC. Also studies have shown that if adenomas are removed then the rate of colon cancer declines (22-24).
4. Pathological investigations show it is very unusual to find a focus of cancer in otherwise normal colonic mucosa, secondly there is often a degree of adenomatous tissue residing in a CRC (19).
5. It is well-known that genetic abnormalities play a key role in cancer development. Support from molecular genetic evidence has mapped the series of genetic alterations in the evolution of adenoma to carcinoma. In normal circumstances oncogenes stimulate cell growth, however

uncontrolled cell growth can result if mutation or overexpression of the genes occurs (19). In opposition to oncogenes, tumour suppressor genes work to inhibit cell cycle progression or promote apoptosis: disruption of these genes results in loss of normal inhibitory control. Another important gene type are Deoxyribonucleic acid (DNA) repair genes which control the rate of genetic mutation. Damage to these means mutations can occur in an uncontrolled manner (18). The key genetic abnormalities related to CRC are:

a) Mutation in the K-ras oncogene (25).
b) Inactivation of tumour suppressor genes: The most important of which is thought to be APC (adenomatous polyposis coli). Other genes thought to play a key role include DCC (deleted in colon cancer), TP53, SMAD2 and SMAD 4 (26).

The detection of adenomas is mainly achieved via endoscopy (specifically colonoscopy or flexible sigmoidoscopy) which is considered to be the gold standard (27). Targeted bowel imaging such as Computed Tomography Colonography (CTC) can also detect polyps with accuracy (28). However, this method is limited to detection only as polyps need to be removed using endoscopy. Due to the clear link between adenomas and cancer, there are stringent guidelines both in the UK via the British Society of Gastroenterology (BSG) and internationally, to advise on endoscopic follow up of patients found to have adenomatous polyps on colonoscopy. The larger a single polyp and/or the higher the number of polyps, the more frequently they will be recalled for follow-up colonoscopy, as shown in figure 2.4.

Figure 2.4. The BSG CRC screening and surveillance flowsheet (27).



2.2.1 Colonic polyposis syndromes and Lynch syndrome

Colonic polyposis is the term given to multiple polyps present throughout the colon. They are classified based on histological type and certain characteristic features (29). Polyposis syndromes are thought to contribute to between 5-15% of all CRCs (30). Most are inherited, cancer risk is higher in almost all of these syndromes.

The most common inherited polyposis syndrome is Familial Adenomatous Polyposis (FAP). It is caused by a germline mutation in the APC gene, located on chromosome 5q21-922 (31). It is inherited in an autosomal dominant manner and has between 80-100% penetrance (32). It affects both sexes in equal proportion with a prevalence of 1 in 5000 to 7500 (29). Multiple mutations have been identified. FAP is characterised by the development of hundreds of adenomas in the second and third decade of life; without treatment almost all those affected will develop CRC (33, 34), with average age at developing malignancy of 39. Other extra-colonic features include upper gastrointestinal polyps, thyroid nodules and brain tumours (35). The diagnosis is made using colonoscopy, where multiple polyps are visualised. The treatment involves surgery to remove the entire colon and rectum

(total proctocolectomy). There are several variants of FAP that form syndromes, characterised by the presence of colonic adenomas and specifically defined extracolonic manifestations, for example CHRPE (congenital hypertrophy of the retinal pigment epithelium) (36) and Gardner's syndrome (osteomas, fibromas and sebaceous cysts) (37). Genetic testing of relatives is essential as stringent bowel surveillance is required in all those affected.

MUTYH polyposis results from mutation of the MUTYH gene, which is responsible for repair of genetic mutations. It is recessively inherited and characterised by multiple adenomas but usually fewer than 100 (33). The affected individuals have a 60-70% risk of developing CRC (33).

Hamartomatous polyposis syndromes include Peutz-Jeghers syndrome, juvenile polyposis syndrome, hamartoma tumour syndromes and other rare syndromes (13). Most confer an increased risk of developing CRC (34).

Serrated polyposis syndrome (SPS) is a more recent discovery and is defined by the World Health Organisation (WHO) as either a) at least five histologically diagnosed serrated polyps proximal to the sigmoid colon, with at least two larger than 10mm b) any number of serrated polyps proximal to the sigmoid in a person with a first-degree relative with SPS; or c) more than 20 serrated polyps of any size distributed throughout the colon (38). In general they are more difficult to detect than pedunculated polyps as they lie flat on the colonic mucosa. The previous understanding was that small serrated (flat) polyps were thought to be most likely hyperplastic and therefore run a benign course. However it has been increasingly recognised that CRC can develop from serrated lesions with the 5-year cancer risk reported as high as 7% (38, 39).

The most common cause of inherited CRC is not a polyposis syndrome. Hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome is an autosomal dominant condition caused by germline mutations in one of the DNA mismatch repair (MMR) genes (40). MMR genes function to repair errors occurring during DNA replication. Interruption in this repair process leads to microsatellite instability (MSI) which is a key feature of this syndrome (34). It accounts for 3-5%

of newly diagnosed CRC (19, 40). Sufferers have up to 47% lifetime risk of developing CRC (41), they are also at an increased risk for synchronous (second cancer occurring at or within six months of surgical resection from index cancer) and metachronus (CRC occurring more than 6 months after surgery for index case) CRCs and are predominantly located in the right colon. The average age at CRC development is 45 years (30), with the progression from adenoma to carcinoma occurring much more rapidly in Lynch syndrome (34). Extracolonic manifestations include endometrial cancer and malignancies of the pancreas, breast and cervix. Current UK colonic screening guidelines recommend at least biennial colonoscopy (27) and have been found to decrease mortality (42).

Hereditary CRC syndromes are being increasingly identified due to better awareness, clear screening guidelines as outlined in by the British Society of Gastroenterology and the investigation for MSI by histology departments routinely on identification of a colonic tumour to detect previously unrecognised Lynch syndrome. Timely identification allows early screening and cancer detection both in the colon and in other high risk sites such as the upper GI tract, but also pre-emptive planning for preventative surgery where appropriate.

Following on from the previous sections on pathological and genetic pathways to CRC, the disease will now be discussed in detail.

2.3 Colorectal cancer

2.3.1 Pathology

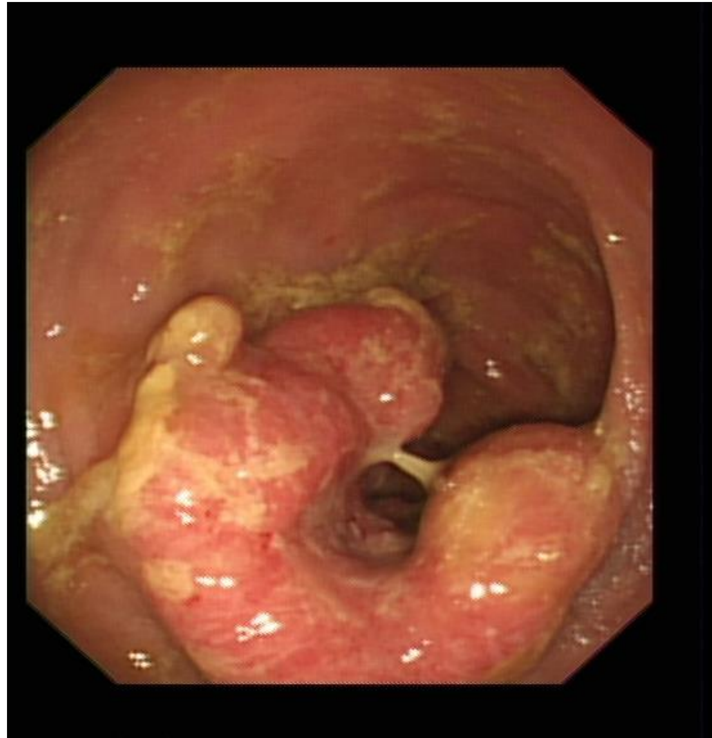
Colorectal (bowel) cancer affects any part of the colon or rectum and cancers are most commonly found distal to the splenic flexure with around 25% affecting the sigmoid colon and 30% affecting the rectum, depending on the patient's sex.

Carcinogenesis is a multistep process, which results in the uncontrolled growth of cells. More than 90% are adenocarcinomas and as discussed above, most of these arise from adenomatous polyps (43) (figure 2.5).

A cancer starts as an intramucosal epithelial lesion. Advancement results in penetration of the muscularis mucosae with subsequent invasion into the lymphatics

and vasculature. The pattern of spread depends on the cancer site. Due to the portal venous system, the liver is the most common site of haematological spread (13).

Figure 2.5. A picture of a colonic adenocarcinoma *detected at colonoscopy (patient permission granted)*



2.3.2 Staging

Two staging systems are widely used to classify affected individuals. In 1929, Cuthbert Dukes proposed a staging classification. Following several modifications, the Astler and Coller classification described in 1954 is most readily used and divides cases into one of three categories: Dukes A –the tumour is limited to the mucosa, Dukes B – Invasion through to muscle layer without lymph nodes involved. Dukes C – Involvement of lymph nodes. Dukes D was added in the 1970s by Whittaker and Goligher and describes the presence of distant metastases (44). The American Joint Committee on Cancer (AJCC) introduced the TNM staging classification for CRC (45), this gives information on the primary tumour (t), the status of regional lymph nodes (n) and the presence or absence of distant metastases (m). Using the TNM information cases are then grouped into one of five stages.

Staging is an important way to categorise cases for prognostication, to guide treatment and to group patients for clinical trial purposes.

2.3.3 Aetiology

Environmental factors appear to play a large role in cancer development. This is evidenced by the regional variations across the world but also variations seen within the UK (46). As mentioned, CRC is predominantly a disease of the elderly and is the result of environmental, dietary and genetic factors coupled with other as yet un-explored susceptibility factors. It is thought sufferers require an accumulation of multiple cell defects with mutational activation of oncogenes together with inactivation of tumour-suppressor genes in order to develop malignancy (43). In this section the factors that have been thought to influence CRC will be discussed.

2.3.3.1 Diet

Diet seems to play a significant role in the development of CRC, as evidenced by higher disease rates in areas where total fat consumption is high such as the UK and the US. Data suggests that obesity is associated with a small increased risk of CRC (47). It is proposed that dietary fat enhances hepatic cholesterol and bile acid synthesis, increasing the amount of sterols in the colon. This increase in sterol metabolites and bile acids seems to damage the colonic mucosa and increase proliferative activity of the epithelium (13). There has been enough evidence of a link between consuming processed red meat and CRC that the WHO's international agency for research on cancer have classed them as Group one carcinogens, albeit with a low risk association (48). Risk seems to be highest with left sided tumours and high cooking temperature is associated with the greatest risk, perhaps due to the production of carcinogens during the charring process. A multitude of evidence has linked dietary fibre to reduced levels of CRC, however, high dietary fibre is not associated with an additional drop in risk (49). It is not clear which fibres hold the most benefit, but it has been suggested that whole grains are particularly beneficial (50). Dietary fibre resists digestion, its protective role is not completely understood but theories include increasing stool bulk, dilution of colonic carcinogens together with promoted elimination which serves to minimise contact with mucosa. Finally,

the fermentation process of fibre by gut bacteria to short chain fatty acids is protective (13, 50). Coffee, fish oil and selenium have also been suggested to provide benefit in CRC development, but this has not yet been substantiated by high quality evidence.

Epidemiological studies have shown a link between dietary fruit and vegetable intake with a risk reduction (51, 52), the relative risk is 0.5 when comparing those with the highest intake to those with the lowest (53). It is postulated that the cancer protection is due to the antioxidant plus other anticarcinogenic components of fruit and vegetables (51).

2.3.3.2 Lifestyle factors

There is a clear link between smoking and alcohol consumption and the development of CRC. Tobacco is associated with an increased risk of and mortality from CRC. Smoking is also a risk factor for all types of polyps (54). Several studies have linked moderate to high alcohol consumption with CRC (55). Alcohol is broken down into acetaldehyde and this has a toxic effect on the colonic epithelium (51). The effect of alcohol on cancer risk seems most pronounced in the distal colon. Finally, physical activity has an inverse relationship with CRC risk.

In summary diet and lifestyle factors hold significance in the development and progression of CRC and our understanding of the mechanisms behind this is growing. Multiple research groups are continuing to investigate these factors and enhance our current understanding further.

2.3.3.3 The microbiome

The human microbiome is the aggregate of the entire community of microbes in the body. There are approximately 100 trillion microbial cells in the human body, which outnumbers human cells ten to one (56). The skin and gut are the primary sites of microbe colonisation. Given the number of microbes present, it is no surprise that the microbiome has a significant effect on human physiology. The body's relationship with it's own microbial community is a delicate balance and alterations or 'dysbiosis' have been linked to diseases such as inflammatory bowel disease (IBD), bile acid malabsorption, diabetes and asthma (57-59).

Several decades ago, it was recognised that CRC not only alters gut flora, but also as a consequence, bodily excretions such as breath, urine and blood. Haines et al published a small study in 1977 measuring the methane concentrations of patients breath with and without CRC and found a much higher proportion with CRC had detectable levels of methane on their breath (60). The surface of the colonic epithelium is exposed to intricate and complex bacterial communities and it has long been suggested that gut microbiota are an important contributor to the health and disease of the gut. Protection occurs by way of preventing overgrowth of pathogens, influencing the development of the innate and adaptive immune system and also provides energy and nutrition to the host.

The identification of key bacteria has developed over the last few years. Understanding of the gut bacterial community comes from stool testing and evaluation of colonic biopsy specimens. Increases in *Escherichia coli* levels have been linked to IBD and also colitis-associated CRC (61, 62) as have reduced levels of *Faecalibacterium prausnitzii* or *Clostridium* clusters; both changes lead to a link between dysbiosis and increased risk of CRC. *Enterococcus faecalis* is thought to release extracellular superoxide which can convert to hydrogen peroxide with subsequent DNA damage and chromosome instability (63, 64).

Ongoing work into this vast microbial community will improve our understanding of the role of bacteria in health and disease, as the primary function of the majority of the gut bacteria remain largely unclear.

2.3.3.4 Inflammatory bowel disease

Inflammatory bowel disease (IBD) encompasses Ulcerative colitis and Crohn's Disease and increases the risk of CRC, accounting for approximately 2% of CRC cases (65). The latest evidence suggests a cancer or dysplasia risk of 7.7% at 20 years and 15.8% at 30 years (27). Risk is linked to duration of colitis and the extent and location of the inflammation, as right sided involvement is associated with a higher risk of developing cancer. There is a well-established endoscopic

surveillance programme in the United Kingdom last updated by the BSG in 2009. This involves a colonoscopy at 8-10 years from diagnosis (27). The interval to next endoscopic assessment is dictated by a combination of endoscopic and histological evidence of inflammation coupled with individual patient risk factors (those with family history or concomitant Primary Sclerosing Cholangitis (PSC) are screened more frequently).

2.3.4 Chemoprevention

Chemoprevention refers to the use of natural or synthetic chemical compounds to reverse, suppress or prevent progression to invasive cancer (66). The results from trials using Aspirin and NSAIDS have shown the most promise. Case-control and cohort studies have demonstrated reductions in incidence of CRC of up to 50%, together with reduction in mortality also observed (67). Although the mechanism is not fully understood, COX-2 has been found to be overexpressed in colonic adenomas and carcinomas. It is thought that these drugs induce apoptosis, and via blockage of cyclo-oxygenase-2 (COX-2) tumour cell growth is inhibited (68). Indeed there has been some more recent trial evidence of a role of COX-2 inhibitors in the reduction of number and size of colonic adenomas; Celecoxib, a COX-2 inhibitor is now approved as an adjunct to standard therapy in patients with FAP (13).

Use of 5- Aminosalicylic acid (5-ASA) therapy has been found to be protective in some studies (69). 5 ASA's are a derivative of aspirin and seem to have both molecular and genetic effects on cancer prevention (70).

Several trials have seen a benefit of calcium supplementation in the prevention of CRC, with a relative risk (RR) 0.80 (71). The American College of Gastroenterology have recommended calcium supplementation for the primary and secondary prevention of colonic adenomas (72). Opinion on the effect of vitamin D supplementation is inconsistent (73).

Folic acid supplementation has gained popularity in the prevention of CRC, but results are conflicting and it may depend if intake is increased by dietary sources or through tablet form. Folate is essential for a number of critical cellular metabolic

pathways, deficiency may contribute to DNA damage, instability of chromatin, impaired DNA repair or aberrant patterns of genomic and/or gene-specific DNA methylation. All these processes can promote carcinogenesis (51). Observational data suggesting a protective effect has not been substantiated by large-scale randomised studies therefore it cannot be recommended as a routine supplement, in fact paradoxically, there has been some evidence that supplementation can be harmful (74).

Other drugs under investigation include the ornithine decarboxylase inhibitor difluoromethylornithine (DFMO) (75)(meyskens), the bile acid ursodiol, epidermal growth factor receptor (EGFR) inhibitors and 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) inhibitors, which are commonly known as statins (76, 77).

The next section will cover the treatment of CRC. Diagnosis will be explored in the subsequent section as it links directly with the review of novel diagnostic markers that features towards the end of this chapter.

2.3.5 Treatment

2.3.5.1 Surgery

Treatment is heavily dependent on stage at diagnosis. Surgical resection is usually the treatment of choice in those with non-metastatic disease. A wide resection of the affected colon is made with removal of lymphatic drainage. Resection margins are dependent on the blood supply and regional lymph node distribution (78). Where possible laparoscopic resection is recommended as it is associated with a faster recovery without any detrimental affect on recurrence or survival (79, 80). Even in patients with metastatic disease it is common practice to remove the primary tumour to prevent complications such as bowel obstruction and bleeding (13).

There is an increasing role for surgery in treating hepatic metastases in carefully selected individuals who have no evidence of extra-hepatic spread. Now 7% of patients with liver metastases are being considered for hepatic resection (81). Less

common is resection of pulmonary metastases, however there has been demonstration of improved survival (82).

2.3.5.2 Chemotherapy

There are a number of chemotherapy regimes, the most convincing evidence of benefit is seen in those with stage III disease. It remains controversial in stage II disease but should be considered in those with high-risk features (81). Drug regimes are frequently superseeded by more efficacious combinations based on new trial evidence. Both adjuvant therapy (chemotherapy used in addition to surgery) and neoadjuvant therapy (given prior to surgery) are both used. Neoadjuvant therapy is commonly used for rectal cancer, but it is not fully clear which of those in the non-rectal CRC group would benefit, therefore the traditional approach would be en-bloc resection followed by adjuvant chemotherapy (81). Chemotherapy has a significant role in metastatic disease too but with modest effects on survival, with one meta-analysis demonstrating 3.7 months additional median survival compared to no treatment (83).

Future advances in the understanding of CRC may allow a more individualised approach to treatment based on anatomical and biological features (84). For example: histological features such as poor-differentiation, signet ring or colloid tumours which are associated with poorer prognosis may be associated with more favourable outcomes in stage ii disease if adjuvant chemotherapy is used.

2.3.5.3 Radiotherapy

Radiotherapy is used to decrease local recurrence of rectal cancers that have penetrated the bowel wall or are associated with regional lymph node involvement. It can also be used for symptomatic relief of rectal bleeding in palliative cases (85).

2.3.5.4 Other therapy

Monoclonal antibodies have been used to disrupt tumour growth and are used in the treatment of advanced CRC. Examples include Bevacizumab that targets VEGF and Cetuximab which targets EGFR (86, 87).

Treatment for CRC is continuously under review and subject to the latest evidence that guides changes. The path to diagnosis of CRC is varied depending on many factors including whether the patient has symptoms or not and encompasses correct utilisation of current biomakers in clinic practice. The next section will discuss the main pathways, and sets the scene for the subsequent discussion on evolving biomarker research.

2.4 Diagnosis of colorectal cancer

CRC's are slow growing and can often go unnoticed by the patient for several years. Asymptomatic patients may have occult blood loss and develop anaemia. Some with anaemia may develop constitutional symptoms such as shortness of breath or fatigue. Common symptoms include change in bowel habit which is more common with left sided tumours, bleeding per rectum (37%) or vague abdominal discomfort (34%) (88). Symptomatic patients tend to have a poorer prognosis. More advanced cases may present as emergencies with obstructive symptoms, such as abdominal pain, distention, vomiting and absence of bowel movements as the tumour affects a large proportion of the bowel lumen. Approximately 25 % of patients with CRC present as an emergency (3). All symptoms associated with CRC can be found in other, more common bowel conditions, making the diagnosis notoriously difficult based on clinical symptoms alone.

Endoscopic investigation of the bowel with flexible sigmoidoscopy or colonoscopy is generally accepted as the test of choice for CRC. It is the only means of obtaining a tissue diagnosis through biopsy and can also diagnose concurrent adenomas, with the potential for them to be removed during the procedure. Imaging of CRC's has progressed over the last few decades, a Barium enema which was once used routinely in the investigation of bowel symptoms has almost completely been superseded by the more sensitive CT Colonography (CTC) (89). In contrast to a traditional CT, the Colonography provides an endo-luminal perspective of the colon and is highly accurate with CRC detection rates over 95% (90) and polyp detection rate is also very good. Patient acceptability for CTC is high (91) and it can be used as an alternative to colonoscopy in more frail, elderly patients. However full bowel

preparation is still required, there is radiation exposure and there is a chance of requiring an endoscopy following this test if an abnormality is found.

Laboratory tests are a useful adjunct in assessing for anaemia and liver function where metastatic disease may be suspected, but have little diagnostic ability in isolation and lack sensitivity (13).

2.4.1 Tumour and biomarkers

According to the National Institute of Health, 'a biomarker is a biological molecule found in blood, bodily fluids or tissue that is a sign of a normal or abnormal process, or of a condition or disease' (92). Biomarkers are being increasingly used in the diagnosis and monitoring of disease. Although biomarkers generally refer to DNA, ribonucleic acid (RNA), microRNA epigenetic changes or antibodies, tumour markers can be considered as biomarkers attributed to the development of carcinogenesis

Tumour-associated antigens (TAAs) are molecules produced by both normal and neoplastic cells but at much higher levels in tumour cells. Carcinoembryonic Antigen (CEA) is the most well-known tumour marker associated with CRC. It was discovered over 50 years ago and is still the only tumour marker recognised to have efficacy in monitoring patients with CRC. It is an oncofoetal protein that rises in the serum of patients with a number of cancers such as pancreas, stomach, lung, breast and bladder. It has no role in the diagnosis of CRC due to low sensitivity and specificity (93, 94), sensitivity ranges between 32% and 69% and is highly dependent on tumour stage (13). CEA is useful in the monitoring of patients with metastatic disease. Rising levels are likely to indicate disease progression and should prompt re-evaluation and possible treatment change (94). CEA monitoring following curative resection has been shown to detect recurrent disease earlier than would be found at routine follow up (95, 96).

Cancer antigen 19-9 (CA 19-9) is a glycoprotein found in serum that currently has no role in the monitoring of CRC, it has inferior sensitivity to CEA and there is no published literature to support its use in any aspect of CRC diagnosis or surveillance (94, 97). Many other protein biomarkers have been and are currently under

investigation for their use in CRC diagnosis, but none are recommended for current use, these are discussed in more detail in section 2.7.1. CA-72 also has poor sensitivity and currently has no use in clinical practice in isolation (97). It may be in future a combination of tumour antigens can be used to provide a tumour antigen profile for the patient which may improve diagnostic accuracy, but this is still work in progress (97).

Faecal calprotectin (FCP) is excreted in excess into the intestinal lumen when inflammation is present (98). Its use in IBD is now well established, as evidenced by its role in both European and American IBD guidelines but more recently there has been interest in the potential clinical application of FCP as a diagnostic adjunct in other pathologies of the GI tract such as microscopic colitis (99), but most prominently in CRC. FCP levels have been found to be elevated in neoplasia, yet a meta-analysis found no significant difference in FCP levels between CRC and controls (100).

Tests involving DNA, RNA and protein biomarkers in stool and blood are being investigated. Biomarker stool tests work on the principle that colorectal neoplasms shed surface cells into the stool. DNA can then be isolated from these cells and tested for mutations and genetic alternations that occur during carcinogenesis (101). For example, detection of methylated SEPT9 DNA in blood yielded a CRC sensitivity of 48% (102), although this does not compete with the sensitivity of the FIT or even the FOBT, ongoing research means newer blood and stool biomarkers with better sensitivity and specificity are likely to emerge in the future (103) and work in this area forms the basis of the majority of experimental work in this thesis.

Microsatellite instability (MSI) describes the changes in coding and non-coding of repeated DNA sequences. These sequences can be exposed to errors in the mutation repair system resulting in the loss of, or multiplication of nucleotide sequence repetitions. Mismatch Repair Genes (MMR) serve to eliminate mutations causing these sequencing errors. Microsatellite instability can be classed as high (MSI-H) or low (MSI-L). Studies have demonstrated a protective effect if MSI-H was present in tumour cells in CRC stage ii and iii with improved survival (104-106). However results when looking at MSI status and chemotherapy regime have demonstrated

mixed results, with studies showing both a better and worse response to 5-FU based chemotherapy with MSI-H (107, 108). As mentioned previously, following analysis of clinical and cost effectiveness, in 2017 NICE recommended universal testing for MMR status for all colonic tumours with the purpose of detecting Lynch syndrome (109).

Loss of heterozygosity on chromosome 18q (18qLOH) in the coding place of SMAD 4 proteins specific to CRC has been a factor of interest as a prognostic marker particularly in CRC stage ii and iii. A systematic review by Popat has demonstrated poorer survival in those with 18q chromosome deletion(110). Research groups have also looked at the role of 18qLOH in predicting response to chemotherapy (111), but there has not been strong enough evidence to date, to base therapeutic decisions on.

Mutation of the p53 tumour suppressor gene is present in 50-70% of all CRC and is associated with a worse outcome. An international collaborative study found p53 mutations present in 34% of proximal tumours, and 45% of all distal and rectal cancers with overall worse survival (112).

Several biomarkers have demonstrated promise when used to predict efficacy to certain treatments. For example the KRAS mutation on the short arm of chromosome 12 is associated with non-response to Cetuximab and Panitumumab. Uridine diphosphate glucuronosyl transferase 1A (UGT1A1) gene polymorphism is associated with higher Irinotecan toxicity and knowledge of its presence can aid in dosing decisions (113, 114).

CRC is a heterogeneous disease with survival outcomes heavily influenced by cell type and stage at diagnosis. It is hoped in the future, with greater understanding of the genetics and biomarker profile of CRC, that treatment can be personalised based on a patient's own tumour profile using genomic and proteomic approaches. The expanding research into biomarkers in CRC diagnosis, prognosis and treatment is an exciting development into the understanding of CRC and will hopefully lead to more individualised therapy with improved outcomes where a prognostic panel of biomarkers are available. Tumour phenotyping with chemotherapy matched

accordingly is perhaps the next step in therapy. Bowel cancer screening using faecal tests will be discussed in detail below.

2.5 Diagnostic pathways in colorectal cancer

Two strategies are widely used to improve CRC prognosis and to optimise the health resources consumed: Population-based CRC screening programs and early diagnosis pathways in symptomatic patients. The criteria for these strategies vary between countries, here in the UK the current system involves:

- Asymptomatic patients over 60 years of age are entering into the bowel cancer screening programme.
- Symptomatic patients of any age meeting referral criteria, being seen via the two-week wait (TWW) scheme.

As discussed above, patients with a risk greater than that of the general population, such as those with IBD, polyposis syndromes and strong family history of CRC are entered into targeted surveillance programmes.

2.5.1 Bowel cancer screening

In 1968 Wilson and Junger's seminal work on screening principles devised criteria for mass screening programs (115). These ten principles, 50 years later, still form the basis of preventative medicine and are followed by the WHO. In essence, for a disease to benefit from screening, it should be important, be detectable at an early stage using a method that is acceptable to patients and have an effective treatment. The screening test should be both sensitive (have the ability to correctly identify those with the disease) and specific (have the ability to correctly identify those that don't have the disease). The natural history of CRC makes it a suitable target for population based screening; it is a relatively common disease with significant mortality associated with it (116, 117) . As discussed in detail above, in most cases it arises from a benign polyp which takes several years to undergo neoplastic transformation. This window of opportunity allows for early detection which is directly associated with improved survival.

Screening for CRC was proposed decades ago, but took subsequent years of robust research before it was rolled out into clinical practice. Several, large randomised controlled trials demonstrated mortality reduction using CRC faecal occult blood (FOB) screening by between 15 and 33% (118-121). Although the technique differed between studies, all demonstrated that screening was better that no screening. Two subsequent systematic reviews have found a pooled mortality reduction in the region of 16% (122, 123). All of these studies used screening based on the detection of blood in stool. Screening for CRC is offered in many countries all over the world but protocol differs between countries and even within the UK. Two UK pilots, one in England and one in Scotland were carried out between 2000 and 2002 to determine the feasibility of CRC screening in the UK population and confirmed the acceptability of FOBT for population screening with response rates of around 60% (124).

Following a funding agreement by the Department of Health, from 2007 the bowel cancer screening programme was implemented nationally. As 80% of bowel cancer cases affect those over 60 year of age this was the age selected for initiation of screening. Initially the upper limit of screening age was 69, this has since been extended to 74 (125).

2.5.1.1 The screening process

In England, from the age of 60, subjects identified for screening are sent an invitation letter. A second letter containing the FOB kit is then sent out a week or two later. Each kit contains six windows and patients collect a stool sample every day for 3 days, using two windows for each sample (figure 2.6) (126). Guaiac FOBTs detect the haem moiety of haemoglobin molecules by making use of the pseudoperoxidase activity of haem; haem will be present if a stool sample contains blood. When a developing agent (hydrogen peroxide) is applied to the stool sample, haem releases oxygen from the hydrogen peroxide which then reacts with the colourless guaiac to form a blue dye. The number of sample windows that turn blue (positive result) is recorded by a technician and further action is determined by the number of positive results. Screening is offered biennially. Kits are sent and received at one of five testing laboratories, called hubs, situated throughout

England. Depending on the outcome of test one, the patient may have a repeat screen, a colonoscopy, or just continue along the screening pathway and repeat the test in two years (127). Those who have a positive test on the first or indeterminate second test will be referred to their local screening hospital to see a specialist screening practitioner within 14 days. The vast majority of patients will be offered a colonoscopy, in less than three percent of cases they will be deemed unfit for colonoscopy and a CT colonography may be offered (126).

***Figure 2.6. A picture of the guaiac faecal occult blood test** with one panel open showing two windows.*

Colonoscopies are only undertaken at centres accredited by the Joint Advisory Group (JAG) on GI endoscopy. Screening colonoscopists undertake dedicated training culminating in a formal assessment. If the patient has a normal colonoscopy, they enter back into the standard screening programme and will complete another FOBT in two years time provided they are still within screening age range. Those with adenomas that are deemed intermediate or high risk are entered into a screening surveillance programme until they reach the age where screening stops, those with low risk adenomas are not surveyed within the screening programme. Those with cancer are managed via the hospital multi-disciplinary team. This process is standardised across the country and comprehensive data collection regarding all aspects (e.g. demographics, uptake of test and colonoscopy,

diagnosis, outcome) is collected for the bowel cancer screening database which is audited at regular intervals.

2.5.1.2 The drawbacks of using the guaiac FOBT

As already discussed CRC screening reduces mortality but the current UK FOBT screening has several drawbacks which affect overall sensitivity and specificity of screening:

1. The guaiac test method is not specific to human haemoglobin, therefore consumption of meat can mean animal haemoglobin is detected in the faeces and will result in a false positive result. Also fruits and vegetables high in peroxidase activity may also create false positives (129).
2. The current technique relies on a technician to detect positivity therefore human error can occur resulting in both erroneously false negative and positive tests (129).
3. Faecal testing is unpopular with patients, and this is reflected in generally low uptake rates. Uptake is around 60%, although this varies markedly between regions (130) with lower uptake in more deprived areas and those with a higher proportion of ethnic minorities. The requirement for three different stool samples on different days is inconvenient for patients and can deter uptake (129, 131).
4. Due to the relatively low sensitivity and specificity, a large proportion of patients unnecessarily undergo a colonoscopy, which is considered an invasive test (129).

Around 2% of FOBT's are positive and approximately 10% of those with a positive test will have cancer. Cancer plus intermediate or high risk adenoma pick up rate collectively has been reported at around 40% (126). Latest regional figures for Coventry and Warwickshire are an 8.8% CRC rate and 50.8% adenoma detection rate in 2017. This compares with the 2014 national figures of 9.4% cancer detection and 59.9% (this includes low risk polyps).

Since the introduction of CRC screening in the UK, two major developments have occurred in an attempt to improve screening techniques and ultimately diagnose more cancer.

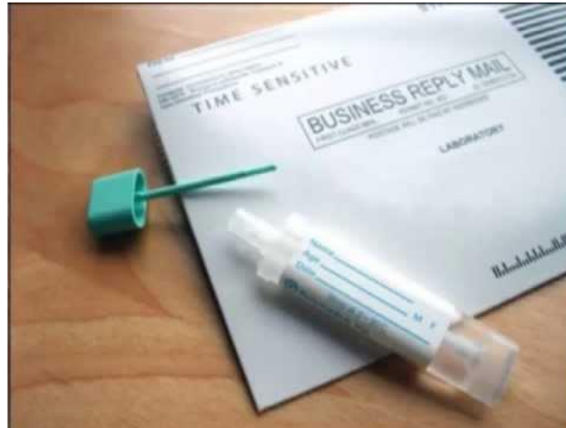
2.5.1.3 Faecal immunochemical testing

In January 2016 the UK National Screening Committee (NSC) published a recommendation that FIT should be implemented to replace the current FOBT. This was decided based on published evidence that FIT:

a) Was easier to use and more reliable given that it does not require human detection and can be measured using machinery
b) Demonstrated increased sensitivity to a much smaller amount of blood than FOBT
c) Is superior to FOBT at detecting advanced adenomas
d) FIT requires a single faecal sample and showed increased participation rates
e) Is as cost effective as FOBT

Implementation of the FIT, which is an immunoassay specific to human haemoglobin within the next few years, will provide a one-stop stool test which has been demonstrated to improve uptake by as much as 25% (126, 130). Immunochemical FOBTs are said to be more sensitive because they use monoclonal or polyclonal antibodies raised against the globin moiety of human haemoglobin, detecting intact human haemoglobin or its very early degradation products. When globin antigen is present, labelled antibodies attach to it, equating to a positive test. FIT collection involves a single sample collected using a wand which is inserted into the stool and then placed in a container (figure 2.7). The process is simpler than the traditional FOBT suggesting that ease of completion is a major contributor to return rates (132). FIT is a quantitative test, the cut off value for screening differs between the countries where it is currently being used as a screening tool. It does allow cut-off levels to be tailored to available resources and colonoscopy capacity; a lower cut-off increases the detection of advanced neoplasia but in turn lowers the positive predictive value and thus results in more demand for colonoscopies.

Figure 2.7. A picture of the faecal immunochemical test kit: *this involves a single sample with applicator within device lid, allowing easier sampling (used with permission from Steve Smith).*



Between 2009 and 2011 the FIT was piloted in the NHS to 66,000 people in Scotland, return rates were significantly higher than for the FOBT (58-61% vs 52-56%) (133). A recent study comparing FOBT with FIT found a 20% higher FIT completion rate, with particular improvement in uptake amongst the lower socio-economic groups(131). However, as this test still relies on faecal sampling, lack of completion by up to 40% of the eligible population means a large proportion remain unscreened (134, 135). FIT is the CRC screening method in many European Countries and the aim is to roll this out to the UK within the next 12 months.

2.5.1.4 Endoscopic population screening

A large randomised controlled trial demonstrated the effectiveness of a one-off flexible sigmoidoscopy in reducing CRC incidence by 23% and mortality by 31% (136). The authors estimated that 489 patients would need to be screened to prevent one death from CRC. Further studies in Italy (137), Norway (138, 139) and the United States (140) confirmed a reduction in incidence and mortality from CRC. The hypothesis that flexible sigmoidoscopy could reduce cancer incidence and mortality is based on observations suggesting that most people who develop a distal colon cancer will have developed an adenoma by 60 years of age (141). The United

States have been using endoscopic evaluation as a screening method for CRC for some time, with consensus that the accrued cost is offset by the reduction in morbidity and mortality from CRC. As discussed earlier in this chapter, distal (below splenic flexure) CRC makes up the majority of CRC seen in the UK (3).

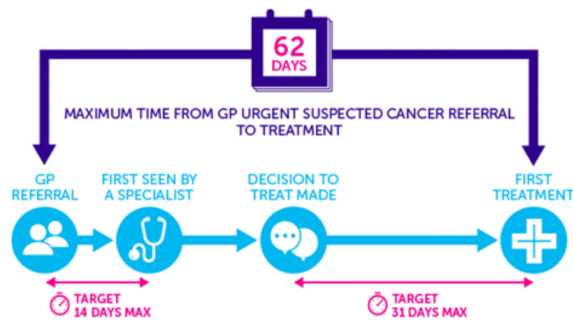
In the UK, since 2013 bowel scope screening has been piloted and from 2015 it has been offered at around two-thirds of screening centres. Currently a one-off flexible sigmoidoscopy is offered to men and women at age 55 in England, this age differs in other parts of the UK. Current randomised controlled trials assessing screening using colonoscopy are ongoing and may well change the face of screening again in the future; results are expected between 2025 and 2034 (142, 143).

2.5.2 The two-week wait (TWW) referral system

Following recognition by the Government and health bodies that patients were facing unacceptable waiting times for assessment, diagnosis and treatment of cancer, a new rapid access strategy was introduced through the Department of Health NHS Cancer plan in 2000 (144). The TWW referral pathway was devised to streamline referral for those with symptoms suggestive of cancer in order to allow diagnosis at an earlier stage, reduce cancer survival inequality around the country, and, ultimately reduce cancer related mortality(144).

By 2001 the Government pledged that those with suspected breast cancer would wait no more than 32 days from referral to diagnosis and this would extend to all cancer by 2005. Targets then focussed not only on diagnosis but treatment, meaning a wait no longer than 62 days from referral to cancer treatment as shown in Figure 11. The government plan to reduce the diagnostic timeframe from 32 to 28 days by 2020(6).

Figure 2.8. A flow diagram of the TWW pathway (145))



Patients with symptoms suggestive of cancer predominantly present to primary care, but although cancer is common, in the primary care setting it is still infrequently seen: for example, a GP working in an average size practice with a list size of around 2000 patients will only see eight or nine new cancers per year, and only one CRC (146) . Thus in collaboration with The National Institute for Health and Care Excellence (NICE), a set of referral criteria to enable standardisation between practices were devised and rolled out alongside the Government referral targets for each suspected cancer site. The criteria include a combination of signs, symptoms and laboratory tests to be used in primary care which have been regularly updated with the latest version published in 2017(147). Patients fulfilling any of these criteria are eligible to be seen by a specialist within 14 days of referral.

The decisions regarding which symptoms should be included, were based on a risk threshold, correlating to positive predictive value (PPV) i.e., if the risk of symptoms being caused by the specific cancer was above a pre-defined level then further investigations are indicated (147). The current NICE guidelines are based on PPV threshold of 3% for symptoms being caused by cancer and this has decreased from the previous guideline from 5%. Recommendations are made based on the best available evidence, prediction models and collaborator expert opinion. In the recently updated guidance, there has been new emphasis placed on faecal occult blood testing using FIT. Although the context of its use is likely to need further clarification as currently ‘unexplained symptoms’ are an indication to carry out FIT testing and if positive then a TWW referral is required (147).

There has been little documentation on specific cancer diagnostic targets via the TWW or any benchmark for numbers of referrals made by primary care trusts. According to the Department of Health (DOH) the original target was to 'identify up to 90% of patients with bowel cancer'(144). This statement now only features in an archived DOH document and does not feature in the most recent publication on the strategy for cancer (5).

The suspected cancer TWW pathway is subject to nine operational standards which detail waiting time measures along with corresponding percentage targets for compliance. Performance data is regularly collected at NHS trusts and national data is available for public viewing. The 2016-2017 annual report shows eight of the nine targets have been met nationally (148). For example, at least 93% of patients referred via this route must be seen within 14 days of referral. According to recent DOH figures, compliance is high with over 95% of patients being seen within 2 weeks (6).

This pathway is not without controversy. A recent study gauging clinician opinion demonstrated significant concern over the disproportionate utilisation of services due to the TWW system, often to the detriment of patients who do not fit this criteria, but make up a greater proportion of total referrals (149).

In summary, methods of improving CRC detection in both the symptomatic and asymptomatic population are in constant development with aims to improve detection at an earlier and more treatable stage. There could be huge benefit gained from devising more efficacious ways to screen and detect CRC via non-invasive methods and without such reliance on symptom-based prediction models. A large proportion of experimental work in this thesis focusses on novel biomarkers in CRC using Volatile Organic Compound (VOC) analysis and are discussed in the next section.

2.6 Volatile Organic Compounds (VOCs)

VOCs are organic chemicals that have a high vapour pressure at room temperature, i.e. that evaporate or sublimate readily under ambient conditions. They have a role in both health and disease in the human body.

2.6.1 The origin of VOCs

VOCs are present in every environment: where people live, work and play. For example, householders may have seen the VOC rating of their paint which gives an indication of the degree of noxious fumes they may be exposed to. The detection and identification of VOCs has been well established in certain areas of industry for years. In waste disposal sites VOC data gives vital information about pollution, providing health and safety information for workers and nearby residents, where exposure to noxious odours is possible (150, 151). Animals utilise VOC information to detect food sources and to distinguish enemies.

Odours omitted from human bodies give information about the metabolic and psychological status of an individual and are recognised by other humans, often without thought; for example, starvation stimulates the production of ketones that can be smelt on the breath and stressed individuals can often emit odourous perspiration. Healthy individuals emit a wide array of VOCs daily depending on their diet, medication, genetics and physiological status (152, 153). This can be considered as the 'odour-fingerprint'. It is not just the more obvious waste products such as exhaled breath and sweat that emanate VOCs, blood is an important source of body odours as many VOCs are derived from blood and are eventually emitted into the external environment via a variety of sources such as breath and urine (154).

The concept of smelling disease was documented by Hippocrates in around 400 BC when he talked of the odour associated with liver disease 'the fetor hepaticus', and many other medical practitioners since have appreciated that disease could change the odour of bodily waste (155, 156). Odour recognition formed the basis of Chinese traditional medicine hundreds of year ago. There has also been significant interest in the use of canines for smelling disease (157, 158). The concept of dogs smelling disease seems logical considering dogs have at least twenty times as many olfactory receptors as humans (159). In 1989, Williams and Pembroke detailed the case of a patient who grew concerned about a mole on her leg due to her dog's persistent interest in it. Indeed the mole turned out to be a malignant melanoma and it was postulated the dog was able to detect the unique smell exuded from the malignant tissue. Even now it is common to see news articles featuring dogs that

have detected disease in small case-control studies (160). However the feasibility of utilising this animal technology as a point of care test in clinical practice is unlikely.

It is only in the last few decades, since the 1980s, that advances in the knowledge of metabolomics and VOCs, together with improved analytical technology, has the concept of smelling disease been studied in more depth using a range of diagnostic techniques (161-163). Consequently, there has been an explosion of interest in researching the use of analytical technology to electronically smell a unique odour exuded from bodily waste products, in order to detect disease by mimicking human olfaction (164-167).

VOC gas phase biomarkers emitted from multiple bodily products (breath, urine, stool, sweat) are an attractive prospect as potential biomarkers of disease. They reflect metabolic processes arising from complex interactions between the gut lining, microflora and invading pathogens (the *fermentome*). These signatures are altered in disease states (154). But metabolic profiling is complex, as VOC signatures found in blood reflect cellular processes in the liver, kidneys, muscles, fat and airways, including bacterial and inflammatory processes (168). Therefore teasing out the disease-specific markers is a big challenge. In order to understand the analysis methodologies for detection of VOCs, a summary of the various techniques is provided below.

2.6.2 Instruments used in VOC analysis

Despite the wide range of studies involving VOC analysis which are discussed below, and the variety of analytical techniques employed, there is no real consensus amongst researchers on which are superior. The decisions regarding instrument choice are usually based on practical factors such as availability and cost together with prior experience. An overview of the commonly used devices is provided here.

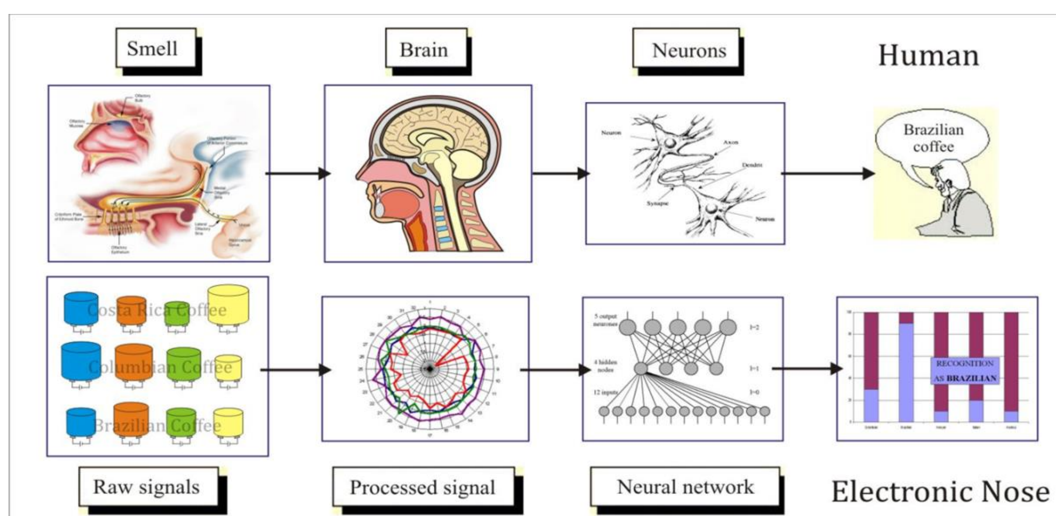
Field Asymmetrical Ion Mobility Spectrometry (FAIMS)
The FAIMS is discussed in detail in Chapter Four. This technology was originally developed for the detection and identification of chemical warfare agents in spiked

food products (169), but has been applied to the detection of disease in several recent studies (170, 171). It is a highly sensitive method of detecting ions based on their mobility, but cannot identify individual chemical compounds.

Electronic nose (e-nose)

First developed in the 1980s in Warwick (172) in an attempt to replicate the human olfactory system, the electronic nose is an umbrella term describing an instrument formed from an array of sensors with overlapping sensitivity. It has the ability to identify a sample based on the composite chemicals triggering sensors within the e-nose (akin to olfactory receptors in the nose). It is composed of chemical sensors that are tuned to different chemical groups with sensitivities in the parts-per-million to parts-per-billion. Figure 2.9 shows the process of sample detection by the electronic nose and how this compares to the human olfaction system. As with the human nose, the electronic nose has the ability to learn through repeated exposure to the same chemical composition (smell), combining these elements to distinguish between different types of sample (170). Samples can be analysed quickly but the e-nose does not have the ability to identify specific chemicals, rather it can distinguish between different classes of samples with distinctive patterns. GC-MS analysis would be required to identify specific chemicals.

Figure 2.9. Schematic of the comparison between the human and electronic nose *(reproduced by permission from Professor James Covington)*

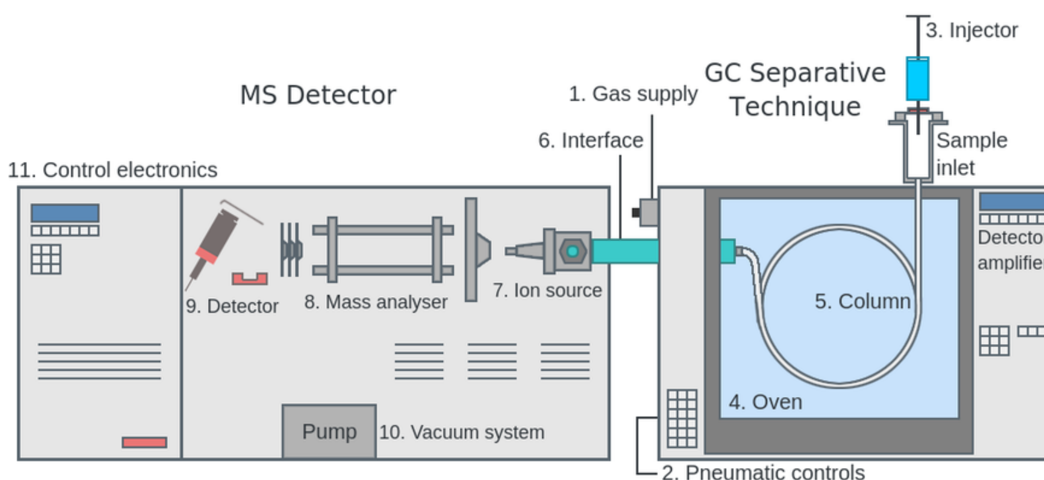


Gas Chromatography – Mass Spectrometry (GC-MS)

This method of VOC analysis can be broken down into its two constituent parts. Gas Chromatography (GC) is a method used to separate the constituent molecules within the sample, according to a combination of their molecular weight and polarity (173). Using long, thin columns with a retentive coating along the inside wall, the sample is passed through a column using an insert carrier gas such as Helium. Gases pass through the column at a mobility determined by the polarity and weight of the molecules within the sample. The column temperature is gradually increased allowing chemicals with lower mobility to pass through. Once chemicals have passed through the column they hit a detector producing a chromatograph based on molecular count and time. This technology can be combined with mass spectrometry (MS).

MS operates by first ionising and breaking the molecules into constituent ion species, they are then accelerated into a right angle bend. At the bend, an electromagnetic field is applied which determines the path of the ions depending on their molecular weight. The ionic mass within the spectrometer when the magnetic deflection is applied, is almost completely unique to an individual chemical compound, a schematic is shown in figure 2.10. GC-MS allows the identity of gases and compounds present within a sample (173). Libraries of compounds and their corresponding mass spectra have been built up and sold commercially by analytical chemical companies. One such library is the National Institute of Standards and Technology (NIST).

***Figure 2.10: Schematic of a GC-MS instrument**. 1-5: GC separates the analytes. The column is flooded with inert gas and the analytes travel along it with speeds dependent on their physical and chemical properties. 7-11: MS is used to identify and quantify the separated substances. Ionisation occurs causing analyte fragmentation. Masses are analysed and the chemical formula can be determined (128).*



Gas- capillary column ion mobility spectrometry (GC-IMS)

This method has been mainly used in the food industry for freshness quality control and authentication of ingredients but it has recently gained interest as a method of disease detection albeit in only a handful of cases to date (174, 175). It is a two-stage analysis of compounds: firstly a gas column provides initial separation, and secondly ionisation of the sample occurs and detection of ions based on their individual mobility as they pass through a drift tube, a further explanation can be found in Chapter Four. Information collated from both steps results in a highly sensitive analytical modality. Its portability, ease of use and low running costs make it attractive for use in the detection of disease. A recent pilot study found 60% sensitivity and 80% specificity in detecting bacterial respiratory tract infections using breath capture (176), but sample numbers were small and more work needs to be done to explore the feasibility of this technique further.

There are multiple other analytical instruments used by different research groups, discussion in this review has been limited to those used in experimental work in this

thesis and an overview of the more frequently used alternatives. The choice of instrument used is a balance between affordability, access and experience in use. Currently, there is no consensus or guideline as to which holds superiority

2.6.3 VOC detection in non-gastrointestinal disease

There is a wide breadth of research into detection of diseases outside of the gi tract using VOCs, these are summarised according to the biological medium used.
Breath

Exhaled breath contains hundreds of VOCs which are derived from both exogenous and endogenous sources. Exogenous volatiles are those that are inhaled from the external environment, those produced following the ingestion of food and those from recreational activities such as smoking, chewing gum and teeth-brushing. Endogenous volatiles are blood-borne, stored in fat and released via the lungs. The basis of disease detection rests on the principle that endogenous VOCs reflect the condition of the cells at the site of disease. As the compounds detected in breath are present in tiny concentrations it can be difficult to distinguish endogenous products from contaminants in the external environment. However, breath collection is very tolerable for the patient and of course is always readily available.

The utilisation of VOCs for disease detection in human breath was described in the 1971 by Pauling and colleagues where a chromatogram of exhaled breath was created (177). Little development occurred for several years after this, mainly hindered by the lack of technical progress of analytical methods until the 1980s.

Some of the earliest and largest volumes of work on VOCs in the detection of disease have focussed on the respiratory system. The majority of work in the respiratory field has been carried out on breath capture, not only using sensor detection but also in isolating chemicals of interest. Breath is mainly made up of nitrogen, oxygen, carbon dioxide and water. The remaining small fraction will consist of tiny concentrations of compounds such as ethane, pentane, acetone and isoprene (178). Most recently in 2018 Pizzini et al isolated four VOCs which were able to distinguish acute exacerbations of COPD from chronic disease with an Area

Under the Curve (AUC) of 0.92 (161) . Separation has also been demonstrated when looking at COPD and controls, with significant differences in hexanal aldehyde concentrations between groups. This may represent oxidative stress in the airways precipitated by smoking (179). Besa et al also isolated six key volatiles which were able to discriminate COPD cases from healthy controls with 70% accuracy, body mass index (BMI), C-reactive protein (CRP) and forced expiratory volume (FEV1) all seemed to affect the breathprint though (180). Despite the number of observational studies conducted with promising results, a systematic review of COPD breath analysis revealed heterogeneity amongst studies with no marker present in all 12 studies included (181).

Other groups have also showed separation in non-malignant respiratory disease including asthma. Using an electronic-nose, asthma has been distinguished from controls (182), and from COPD with an accuracy of >90% (183) but separation of mild from severe asthmatics was not demonstrated (182). Tuberculosis has also been studied with success using e-nose: 100% accuracy was reported by Pavlou et al (184) and with a sensitivity of 84% by Van de Schee and colleagues(185). Studies involving mesothelioma, cystic fibrosis and sarcoidosis have also been published (186).

In malignant disease, O'Neill and colleagues identified 28 breath VOCs as potential markers of lung cancer (187). Similar results were found by Phillips et al who identified 22 potential VOCs associated with lung cancer. These were predominantly alkanes, alkane derivatives and benzene derivatives with 100% sensitivity and 81.3% specificity for cancer detection (188). Schallschmidt and colleagues discuss the limitations of breath capture with low and variable VOC levels detected at times despite high separation of lung cancer from control, albeit on small patient numbers (189).

Other malignancies have been studied using breath capture, in particular breast and urological cancers have attracted interest (190). A recent study using GC-MS found test accuracy of 90% for detecting breast cancer in 54 patients from 124 controls

(191). Others have demonstrated that the breath VOC profile differs in melanoma, mesothelioma and head and neck cancer: many are pilot or proof-of-concept studies (192-195). To date, hundreds of breath VOCs have been identified with a proposed link to various cancers, however no consensus on those key to identifying malignancy has been made (196). Finally, another area that has gained interest for breath VOC profiling are liver diseases such as hepatocellular carcinoma and cirrhosis (197, 198).

Breath analysis raises a number of important challenges despite some obvious advantages in disease detection. It is not clear what the relationship is between alveolar concentration of a compound and that found in exhaled breath, also the stability of these compounds is limited and repeatability seems to be a particular issue (199). Breath is the most vulnerable detection source to the effects of environmental factors and overcoming this will be the major challenge for this method of VOC capture.

Urine

Urinary metabolites are the end products of many metabolic pathways, the use of urine in the diagnosis of a number of metabolic disorders such as diabetes, carcinoid syndrome and porphyria is well established. Urine contains a complex mix of components that could be significantly influenced by food and drink ingestion, this can prove challenging when trying to separate compounds of interest. However, urine is readily available, collection is usually acceptable to patients and it has the potential to be stored long-term with relative stability. It is the second most investigated modality of VOC detection after breath. Although the focus of urinary analysis has largely been on gastrointestinal disorders, investigation into a diversity of urological disease and other malignancies has also taken place and is summarised below.

Using a model based on four VOCs prostate cancer was separated from control with modest accuracy of 63-65%, this increased to a maximum of 74% when VOC

profile was combined with prostate-specific antigen (PSA) (200). Bernabei also looked at prostate cancer and bladder cancer comparing with healthy controls before and after surgery. The differentiation of both cancers from control was demonstrated pre-surgery, but this was not replicated when post-surgical cancer samples were compared with controls (201). A study of 22 patients with renal cell carcinoma (RCC) were compared with healthy controls using solid phase microextraction (SPME) followed by GC-MS. They found 11 elevated VOC markers and 3 decreased in the RCC group compared with controls (202), but acknowledged these results would need to followed up with a larger study. Detection of urinary tract infections (UTI) has been attempted by five research groups (203), most recently by Roine et al in 2014 where the detection of four common pathogens responsible for UTI were tested against sterile plates (204), sensitivity was 95% and specificity was 97% for detecting the correct bacteria.

In order to investigate a faster method of Tuberculosis detection, a study was conducted to identify urinary V0OCs in those with tuberculosis versus controls. Results showed a significant ($p < 0.05$) increase in the abundance of o-xylene (6.37) and isopropyl acetate (2.07) and decreased level of 3-pentanol (0.59), dimethylstyrene (0.37), and cymol (0.42), the authors felt this information could form the basis of a diagnostic test (205). In pilot work, the FAIMSHED research group has used VOC detection to differentiate control samples from inflammatory bowel disease, liver cirrhosis, diabetes and obesity (154, 206, 207) and most recently pancreatic cancer where pancreatic adenocarcinoma was detected from health controls with 91% sensitivity and 83% specificity with 81 patients in each group (208).

Urine definitely holds promise for VOC diagnostics, perhaps particularly in urological conditions and gastrointestinal given the promising pilot work carried out to date. The ease of collection and storage is certainly attractive, but repeatability of much of the work mentioned has not been demonstrated on a larger scale.

Blood

Blood VOCs are an attractive prospect as they directly reflect the internal environment and are not altered by external environmental factors. There is also the potential to store samples for prolonged periods. However collection is relatively invasive for patients, plus sample centrifuging is time consuming therefore this may seem a less appealing option. Consequently there has not been a large body of research conducted using VOC blood markers. Blood VOCs in the diagnosis of CRC and other GI disease are discussed below.

Faeces

Stool samples in theory may seem the most logical way of detecting VOCs particularly in gastrointestinal disease, as they represent the dietary end-products of digestion and intestinal bacterial metabolism. However the patient's microbiome is known to significantly influence faecal VOC profile and as demonstrated by the modest uptake in bowel screening, patients do not like providing faecal samples. Also samples are unlikely to be produced at the time of a clinic, thus rely on the patient returning the sample at a later date. These factors all prove a barrier in the clinical setting.

Other

Some of the earliest work in measuring aromas was used in diagnosing the presence of microbial pathogens, (209-212) such as detecting beta haemolytic streptococcal infection in leg ulcers (213).

Sweat is the main source of VOCs found on the skin. Although commensal bacteria present on the skin surface, can alter the compounds secreted in sweat thus making analysis difficult. Saliva has also gained interest and may be able to reflect VOCs formed in blood (214). Vaginal secretions have been studied, especially in the diagnosis of pelvic infections as normal vaginal secretions are virtually odourless therefore contamination is minimal. Finally bile VOCs have shown some promise, but only by a single research group, in both the detection of malignant biliary

strictures in pancreatic cancer and cholangiocarcinoma associated with Primary Sclerosing Cholangitis (215).

2.6.4 VOCs in gastrointestinal disease

The diagnosis of a variety of GI pathology has been investigated using breath, urine, faeces and blood. One of the earliest breath studies was in 1992 when Sivertsen et al performed a small study looking at whether breath methane excretion differed between those with CRC and controls: they did not find any association (216). Studies reporting breath analysis in IBD and other GI disorders are lacking. A small body of evidence has linked breath alkanes with colonic inflammation (217), with ethane, propane and pentane significantly elevated compared with healthy controls (218).

IBD has been the most studied disease in faecal headspace analysis with studies looking at IBD versus control with 76% sensitivity and 88% specificity (219). Using GC-MS, Walton and colleagues found a striking increases in ester and alcohol derivatives of short chain fatty acids in Crohn's disease compared with the other groups. They also found that following treatment for Crohn's disease VOC profiles were more similar to healthy controls (220). Active versus inactive IBD has also been evaluated using GC-MS (221). A detailed analysis was published in 2007, in which Garner et al explored the difference in faecal VOC emissions in health and in the presence of gastrointestinal disease. Using GC-MS, they analysed stool samples from those with UC, *Campylobacter jejuni* and *Clostridium difficile*. This was the first detailed analysis of faecal VOC emissions. A total of 297 VOCs were identified, the main classes of volatiles included (222):

1. Acids, alcohols and esters – Short-chain fatty acids were seen commonly and arise form the metabolism of undigested carbohydrate
2. Benzenoid and heterocyclin compounds
3. Aldehydes and ketones – Ethanal which was present in all samples is thought to be associated with bowel cancer
4. Alkanes, alkenes, alicyclic compounds –
5. Ether compounds and chloro compounds

6. Nitrogen and sulphur compounds

This paper confirmed the hugely varied VOC profile of faeces. Some compounds can be explained as breakdown products of certain foods, others will be explained by the alterations in microflora in the presence of disease, as there may be less microbial flora with consequent reduction in secondary metabolites. Other changes were difficult to explain.

Through bacterially mediated gastrointestinal disease, several other potential faecal biomarkers have been detected. *Escherichia coli* has been linked to inflammatory gastrointestinal diseases, biomarkers such as terpenes, trimethylamine and ketones have been isolated from infected stool samples (184, 223). Through faecal odour analysis Moore et al and Suarez discovered that sulphur-containing compounds such as hydrogen sulphide, dimethyl disulphide, methyl disulphide and dimethyl trisulphide contribute most to faecal odour(224-226), these findings contribute to our understanding of the possible background VOCs that may be present in most normal faecal samples. Rossi et al recently assessed the effect of diet on the faecal VOC profile of patients with functional bowel disorders. Here patients were allocated to a low FODMAP (fermentable oligosaccharides, disaccharides, monosaccharides and polyols) diet or a sham diet with or without the addition of a probiotic, using a GC sensor device they found key VOCs associated with response to adopting the low FODMAP diet and response to probiotic (227). This study is the first to raise the possibility of using VOC profiling to predict response to dietary interventions in the functional bowel population. Coeliac disease and radiation toxicity have also been studied using faecal biomarkers (228).

Urinary markers in GI disease have been explored by Arasaradnam et al with IBD vs control study using the FAIMS instrument and the electronic nose. They found that both technologies were able to distinguish IBD cases from controls with >75% accuracy (*p=0.001*), replication of these findings in a larger study has not occurred but would be interesting (171). The same group also compared coeliac disease urinary samples with IBS-D controls, here 1,3,5,7 cyclooctatetraene was considered significant in the coeliac group but not the control (154).

Blood analysis has gained some interest, although sampling is less attractive to patients. In theory blood may be more able to represent the internal environment affected by GI disease and be less influenced by the external environment that can be potentially problematic in breath analysis. Only a handful of studies have explored blood VOC biomarkers. In 2016 Bhatt et al explored oesophageal cancer in a pilot study by looking at the plasma VOC profile of 20 cancer patients compared with 19 with gastro-oesophageal reflux disease (GORD), test accuracy was good with AUC 0.83 using ion flow tube mass spectrometry for analysis. They also identified a potential nine VOCs of interest in the cancer group (229). Hepatic encephalopathy in cirrhosis has been studied with 3-methylbutanal implicated as a potential biomarker, although results have been mixed (230). Using GC-MS, detection of blood biomarkers for liver cancer were explored and found hexanal and 1-octen-3-ol to be of potential interest (231).

2.6.5 VOCs in the detection of colorectal cancer

Most clinically relevant VOCs in CRC are thought largely to consist of benzene, alkanes and aldehydes and their derivatives (232). In cancer, several mechanisms have been postulated to cause alteration in the VOC profile:

- Tumour growth creates gene and protein changes that can cause peroxidation of the cell membrane and subsequent release of VOCs (48).
- There is increased prevalence of reactive oxygen species within a cancer cell causing alteration of the VOCs and induction of cytochrome p-450 enzymes (232-234).
- The microbiome, or human gut bacterial population is known to affect the VOC profile and there is growing evidence supporting the role of infections in cancer (235).

Canine detection of CRC has also been explored by Sonoda et al where the sensitivity of canine scent detection of breath samples compared with colonoscopy was 0.91 and the specificity was 0.99 (158).

A recent useful review by Di Lena et al identified ten papers (164-166, 190, 232, 235-239) examining VOC biomarkers in the detection of CRC (240). Table 2.1 summarises each study.

Table 2.1: Summary table of publications reporting VOC biomarker use in detection of CRC *(up to 2016). Adapted from reference (240)*

First author	Year	VOC source	Analysis	Patient numbers	Sensitivity/specificity
Peng G	2010	Breath	E-nose GC-MS	26 Cancer 22 Control	Sensitivity: <30% Specificity: no data
Silva CL	2011	Urine	SPME/GC-MS	12 Cancer 21 Control	Not stated
Altomare DF	2013	Breath	GC-MS	37 Cancer 41 Control	Sensitivity: 86% Specificity: 83%
Wang C	2014	Breath	GC-MS	20 Cancer 20 Control	Not stated
Arasaradnam RP	2014	Urine	FAIMS/GC-MS	83 Cancer 50 Control	Sensitivity: 88% Specificity: 60%
Wang C	2014	Blood	GC-MS	16 Cancer 20 Control	Not stated
De Meij TG	2014	Faeces	E-nose	40 Cancer 60 Advanced adenomas 57 Control	Sensitivity: 85% (CRC), 62% (adenomas). Specificity: 87% (CRC), 86% (adenomas)
Batty CA	2015	Faeces	SIFT-MS	31 High risk 31 Low risk (based on colonoscopy)	Sensitivity: 72% Specificity: 78%
Westenbrink E	2015	Urine	E-nose	39 Cancer 35 Irritable bowel syndrome 16 Control	Sensitivity: 78% Specificity: 79%
Amal H	2015	Breath	GC-MS E-nose	65 Cancer 12 Advanced adenoma 10 Non-advanced adenoma 122 Control	Sensitivity: 85% (CRC vs controls) Specificity: 94% (CRC vs controls)

Key

E-nose: Electronic nose
GC-MS: Gas Chromatography Mass Spectrometry
SIFT-MS: Selected Ion Flow Tube Mass Spectrometry
Dhs-SPME: Dynamic Solid-Phase Microextraction in headspace mode

Of the papers reported in the Di Lena review, eight studies isolated VOCs that were found to be associated with CRC. Using SIFT-MS, Batty et al tentatively identified hydrogen sulphide as a key VOC (235) with high levels of sulphide purported to cause toxicity and potentially DNA damage at the genomic level. However the drawbacks of faecal analysis have already been discussed. Ammonia and acetaldehyde were also found in higher quantities in the cancer cohort. Only one paper has specifically targeted the screening population by examining the use of faecal VOCs as a screening test for CRC (235). The heterogeneity between study methodology was high and a there was a wide variation in compounds found between the studies with an array of alcohols, alkanes, sulphurs and others meaning little overall conclusion can be made about the key biomarkers. However this could provide the basis for further work into compound detection and also postulations about the origin of these compounds in relation to cancer development. Work into faecal analysis continues, recently a publication showed alterations in faecal microbiota and metabolome in samples from CRC cases versus controls. The abundance of *Fusobacterium* and *Porphyromonas* had a significant effect on the metabolic profile (241).

Three papers focussed on urine: Silva et al performed a study involving three cancer groups (leukaemia, colorectal and lymphoma). For colorectal cases there were 12 cancers and 21 control and analysis with two-stage using SPME followed by GC-MS. A total of 82 volatiles were identified with three felt to be most characteristic of the CRC samples (1,4,5-trimethyl-naphthalene, 2,7-dimethyl-quinoline and 2-methyl-3-phenyl-2-propenal) (164). In 2014 Arasaradnam et al performed a prospective case control study comparing 83 unselected CRC urine samples with 50 controls using the FAIMS (field asymmetric ion mobility spectrometer) instrument. Using Fisher Discriminant Analysis the sensitivity and specificity to detect CRC were 88% and 60% respectively (236). The study reported the first evidence of urinary diagnostics in CRC, postulating that healthy controls can be distinguished from those with CRC using their VOC profile with accuracy. The same research group headed by Westenbrink investigated 39 CRC patients and compared these with IBS and control groups, sensitivity and specificity were

modest at 78% and 79% respectively. Using GC-MS this group identified the following 9 urinary chemicals in sporadic CRC. Chemicals include Acetone, 2-Pentanone, 4-Heptanone, 1,3,,5,7-Cyclooctatetraene, Allyl Isothiocyanate, Oxime-,Methoxy-phenyl, 1,3-Propanediamine, Carvone, Ethanone 1,1’ (1,4-phenylene)bis, Phenol, 2,4-bis (1,1-dimethylethlyl)- (238).

2.6.5.1 VOC profile and cancer stage

It is plausible that changes in cancer stage will alter the VOC profile progressively. The multifactorial effect of cancer on the colonocytes, microbiome and vasculature will all contribute to metabolomic changes. It has been postulated that the abnormal proliferation of cancer cells is key to the metabolic profile changes that differentiate cancer from controls (242). But it is unclear whether these gradual changes can be detected by the current technological abilities in VOC analysis.

Given that cancer stage is key in survival (243), a central component of the work in this thesis is to investigate whether specific VOCs can predict CRC stage. Two publications explore VOC profile and CRC the stage (Table 2.1): Altomare et al used breath analysis to try and identify differences in VOC profile in cancer and control using GC-MS. When categorised by stage they found no difference in patterns when comparing stage I/II with III/IV cancer (232). Numbers were small in this study, with a total of 37 cancer subjects. Wang et al found no separation trend between early and late stage CRC in blood VOC analysis, but just 16 cancer patients were included. In order to answer this question more robustly conducted studies, with larger numbers are needed (237).

To summarise, there is an increasing body of interest in VOC diagnostics but consensus is lacking on methodology and target VOCs. Diagnosis of CRC in particular, has raised interest with some encouraging results demonstrating altered VOC profiles. These have the potential to form the basis of a diagnostic test, although currently there is no VOC based marker in use in clinical practice.

In conjunction with the above review of urinary biomarker diagnostics, the next section will provide the background to the experimental work completed in Chapter Six on urinary peptide analysis. This area of research continues to increase interest

amongst researchers, although little has been carried out to investigate CRC specifically.

2.7 The proteome

The proteome describes the entire protein make-up of an organism, such as a human. Proteins are the building blocks for cells, and provide a messaging role which is active in every cell and tissue in the human body. Analysis of these proteins has established key information on the composition and regulation of cellular pathways (56). Although the genome has been mapped, without understanding the consequent protein changes, knowledge about the pathogenesis of a disease process at the cellular level is limited. As yet mapping of the full human proteome has evaded researchers, but The Human Proteome Project aims to achieve this in the coming years (244). Difficulties in achieving proteome mapping are due to the huge numbers of proteins present in the body. Not only do they outnumber genes considerably, due to alternative splicing and post-translational modifications, there are also multiple copies of each protein that can alter their states, based on the cellular environment (245). Additionally, the dynamic nature of protein expression relates to various human factors including temperature, presence of disease, medications, diet and many others. This makes stabilisation and reproducibility of sampling extremely challenging. Finally, there is no expansion technique akin to the polymerase chain reaction (PCR) in DNA that can currently be applied to proteins. Previous studies have refuted the theory that mRNA levels would correlate with proteomic information, in addition translation does not occur at a constant rate and post-translational processing and degradation occur at varying speeds (245, 246). Despite these difficulties, understanding the huge complexities of the proteome could result in an un-paralleled understanding of disease, and in particular cancer, heralding new targets for therapy and early detection (245).

Proteomics is the study of proteins in a tissue or biological fluid and gained interest around twenty years ago, as a new focus for the identification of therapeutic targets for disease (247). It was postulated that polypeptides excreted in bodily fluids were likely to reflect bodily functions (248). Proteomics aims to complement genomic processes to better understand biological functioning. There are estimated to be

around one million proteins in the human body, and as mentioned above, these proteins are in a constant state of change.

Thousands of research groups have reported on protein-based disease biomarkers. Due to high biological variability it has been concluded by many of these researchers, that evaluation for the presence of a clinical condition on the basis of a panel of biomarkers, rather than a single peptide is far more likely to yield clinically useful results (249-251). This allows for some degree of flux in individual analytes. Also, as carcinogenesis is a highly complex process influenced by vast numbers of elements such as genetics and environment, a single peptide marker is unlikely to be consistently present or always clinically relevant. Various bodily fluids offer a target for protein analysis and it is not clear whether one is superior to the other, one might postulate that sampling close to the site of disease would be logical e.g. breast ductal fluid in breast cancer or pleural fluid for lung disease but results to date have not consistently proved this theory (250).

Renal diseases were the first major area of interest for proteomics in the detection of disease, with urine analysis being favoured partly due to ease of collection and availability and partly due to the hypothesis that it would closely match the system of its origin i.e. the kidney and urinary tract. In 2000 Mckee et al were one of the first groups to recognise the potential of complex protein analysis in urine by observing the presence of integral membrane transporters in the urine of mice (252).

2.7.1 Proteomics and cancer detection

The application of proteomic technology in the discovery of cancer biomarkers have been numerous and would be impossible to summarise. There are several well-known and established single protein biomarkers, mainly detected in the serum (see table 2.3) and used to aid diagnosis, risk stratify and others to provide prognostic information and to aid in deciding therapy.

Table 2.3. Approved purely protein-based biomarkers in cancer

Biomarker	Source	Cancer	Use
CEA	Serum	Colorectal	Monitoring
EGFR	Colonic	Colorectal	Selection of therapy
PSA	Serum	Prostate	Monitoring
HER2	Serum/breast tumour	Breast	Prognosis/selection of therapy and monitoring

CEA: carcinoembryonic antigen, EGFR: epidermal growth factor receptor, PSA: prostate-specific antigen, HER2: human epidermal growth factor receptor 2.

As mentioned above, theoretically, protein derived ‘fingerprints’ based on an array of peptide markers may have the potential to outperform current protein-based biomarkers. Some of the earliest work involving polypeptide expression or the ‘peptidome’ was in breast cancer tissue versus benign fibroadenoma tissue. The investigators found heterogeneity of peptide expression between and within groups but there were some patterns tentatively reported (253). Around the same time Hirano et al studied lung cancer and Paweletz et al correlated the protein annexin 1 with early onset oesophageal and prostate cancer (254, 255). Prostate cancer proteomics have been studied in tissue, serum and urine, in general study numbers have been small but tissue profiling seemed to be the most accurate (247, 256, 257).

There has been some controversy over results involving the serum diagnosis of ovarian cancer using peptide analysis: Petricoin and colleagues were heavily scrutinised and re-analysis of data by a separate group revealed potential bias that may have resulted from change in methodology between analysis and perhaps the authors were holding significance with small peptides usually discounted as environmental or experimental artefacts (258). This incident highlights the drawbacks of this technology and the difficulties in establishing a reliable test

because the peptidome is in constant transition in response to factors such as environment, medications and recreational activities.

In the last two decades there have been multiple studies investigating the role of specific protein effects on the chances of diagnosis, stage at diagnosis, response to treatment, chances of recurrence and many other parameters important in the diagnosis and treatment of malignant disease. Despite this work, investigation into the potential diagnosis of a variety of cancers using protein analysis has been met with mixed success and in most cases their role as a cancer biomarker is as yet not validated for a defined clinical application. Studies reporting predictions of response to treatment have been more encouraging: Okano et al detailed using the proteomic signature to assess likely response to Gefitinib, nine proteins were distinguishable in responders from non-responders (259). In addition, Taguchi and colleagues used pre-treatment serum to predict benefit from an inhibitor of EGFR tyrosine-kinase in those with non-small cell lung cancer. Following use on a training set, they applied this model to validation cohorts, although there seemed to be merit in the predictions the actual proteins were unable to be identified therefore no assay could be developed (260).

CRC has gained a large amount of interest in the field of proteomics, this is, in part, due to its high prevalence in the Western world, overall poor prognosis and lack of sensitive non-invasive tests for the disease. Many studies assessing protein-based biomarker utilisation have focussed on disease diagnostics (261, 262). Possibly the most promising markers to date, although not proteins, are DNA methylation markers based on the number of subjects and published reports (263, 264) . Samples used have been mainly tissue or in some cases serum. Blood based biomarkers have received the most attention in identifying panels of proteins with potential as diagnostic tests. Recently Fung et al used a three protein panel test to differentiate CRC from control with 73% sensitivity and 95% specificity in early stage disease (265). Urinary proteomics has gained some interest with studies focussed on urological cancer but thus far there is a paucity of proteonomic studies involving urinary CRC detection. Understanding the complex biology behind CRC development using proteomics has the potential to add knowledge not just to the

diagnosis of CRC but in the cancer development and to the risk stratification and profiling of patients for targeted therapy.

2.7.2 Instruments used in protein analysis

Proteomic techniques have changed and adapted as understanding of the technology has improved. Also as the complexities of protein and peptide mapping are discovered, the way data is captured has had to evolve alongside this. In the early stages of protein analysis a technique called Western blotting was employed (266), subsequently other immunologic methods of analysis gained popularity, such as immunohistochemistry and immunocytochemistry using enzyme-linked immunosorbent assay (ELISA). But these techniques are limited as they can only identify a small number of proteins at each step. This is helpful when the target proteins are already identified thus narrowing the targets of interest exponentially.

Two-stage analysis was in development from the late 1990s, following refinement of well-established protein electrophoresis methods coined by Klose et al (267). Popular techniques included polyacrylamide gel electrophoresis (PAGE) and two-dimensional electrophoresis (262) for protein separation combined with MS. MS was identified as a developmental breakthrough in the analysis of proteins and has become the central element in most projects today (268). Complementing protein capture with MS was subsequently recognised to be the way forward in isolating protein data in the extremely complex and fluctuant environment of the cell. Although these techniques of proteomic separation can detect thousands of protein spots, MS must be applied to each individual spot in turn, meaning use outside of the research field would be impractical and unsuited to the clinical environment (248). Hence the trade-off for more detailed analysis is time consumption (250). Technological advances have constantly been in development however, with the inefficiencies in the currently used techniques subject to ongoing refinement.

Several other micro-array methods for capturing proteins have since been complemented with MS and are summarised in table 2.4. For example, surface-enhanced laser desorption and ionisation (SELDI) MS is another important tool as it allows binding of specific fractions of polypeptides in a complex sample which

reduces the number of peptides requiring analysis using MS (269, 270). The SELDI-MS technique also has the advantage of speed of sample analysis, with the ability to sample multiple cases in one step without requiring specialist skills for operation (271). The disadvantage is that this only represents a small section of the samples limiting the peptide profile and therefore the diagnostic potential. It is however, still a popular mode of analysis. For example, Petricoin et al analysed serum in the diagnosis of ovarian cancer (272) and Hampel et al evaluated the urinary proteomics before and after radiocontrast administration using SELDI-MS (273). Ward et al also used SELDI-MS in two studies that focussed on serum and urine proteomics in CRC versus control demonstrating high sensitivity, earlier work by Chen and Yu also saw success using this technique in CRC proteomic detection (274, 275).

Capillary electrophoresis developed shortly after SELDI-MS, and is widely accepted as an effective analytical technique with several advantages (276). These include a constant stable flow, thus avoiding a gradient which may change the ionisation parameters. Running costs are low and it also allows rapid separation, uses a small sample size and combined with MS is a powerful tool in proteomics (271). Furthermore it is also compatible with nearly all buffers and has a high resolution separation, resulting in a high number of potential biomarkers displayed (249). Multiple modifications and resultant improvements in technique over the years are outlined in a review article by Dolnik et al (277).

Several groups have demonstrated the diagnostic potential using this technique in the field of cancers including renal, bladder, pancreatic and prostate (249, 278-280). Most recently belczacka et al performed a large study investigating urinary peptide profiles captured using CE-MS in five different cancers, to assess whether a general tumour marker was possible (281).

CE-MS has been used for all proteomic experimental work in this thesis due to the advantages outlined above together with several years of previous experience using this technique by the collaborator group. A schematic demonstrating the steps of CE-MS analysis is shown in Chapter Six.

***Table 2.4 Analytical techniques currently available for proteomics** with advantages and limitations listed (reproduced with kind permission from Mosaiques diagnostics)*

Technology	Advantages	Limitations
2-DE-MS	Applicable to large molecules.	Not applicable to peptides <10 kDa, no automation, time consuming, quantification difficult, expensive.
SELDI-MS	Easy-to-use-system, high through-put, automation, low sample volume required.	Restricted to selected polypeptides, low-resolution MS, low information content.
LC-MS	Automation, multidimensional, high sensitivity.	Time consuming, sensitive towards interfering compounds, restricted mass range.
CE-MS	Automation, high sensitivity, fast, low sample volumes required, multidimensional, low costs.	Not suited for larger polypeptides (>20 kDa).

Key: DE-MS – differential electrochemical mass spectrometry , SELDI-MS – surface enhanced laser desorption/ionisation, LC-MS- liquid chromatography mass spectrometry, CE-MS – capillary electrophoresis mass spectrometry

As analytical techniques for protein detection have improved and continue to be modified, so too have the bio-informatics techniques. This is the process by which raw proteomic data is processed and through machine learning algorithms, (where known outcome data is inputted) a training set of samples can be used to predict the probability of a diagnosis of an unknown sample (269). So-called disease and control clusters are created as a training set model. As new patient data is analysed they are classified as disease or control based on the clusters they fall into. They are also scored based on their proximity to a certain cluster-the higher the score the higher the probability a sample belongs to that cluster group.

Proteomics holds enormous potential, but until control for the sample variables can be improved or at least accounted for in a standardised format, it won't be clear whether this technology has a firm place in future diagnosis of disease. Ongoing developments in proteomic technologies allow increasingly sophisticated ways of identifying the differentially expressed proteins in various bodily fluids, creating a continual wave of excitement and possibility in this area.

2.8 Conclusion

CRC carries a significant disease burden in the Western world and a large amount of work into improving diagnostic methods has been carried out in the UK and across the world, yet survival figures have not significantly changed in recent years. Finding ways to diagnose the disease earlier is key. Not only does a new test have to be suitably accurate but it also has to be acceptable to the patients, hence urine holds enormous promise. Identifying new biomarkers of CRC is an exciting and rapidly evolving area of interest. Extensive review of the existing literature confirms the work in this thesis is the first to systematically investigate the TWW pathway for the detection of CRC. In addition, the experimental work, both on gas-phase and peptide biomarkers provides original results looking at the bowel cancer screening population and those with symptomatic disease.

CHAPTER 3:

The effect of the two-week wait referral system on the detection of colorectal cancer: a systematic review and meta-analysis

3.1 Introduction

Colorectal Cancer (CRC) is a significant cause of death in the UK so establishing ways to improve cancer detection and at an earlier stage to improve survival are vital. As discussed in Chapter One, national survival outcomes rank poorly when compared to many countries in Europe. The two-week-wait (TWW) cancer referral pathway was devised in 2000 to standardise referral and investigations for suspected cancer and ultimately to speed up the diagnostic process. The lower gastrointestinal (GI) referral pathway has been subject to several updated versions since inception, the latest can be viewed on the National Institute for Clinical Excellence (NICE) website (147). Each updated version of the guidelines has resulted in a widening of referral criteria and consequent increase in referral numbers. The TWW pathway is a predominantly symptom-based referral system for a disease that is known to have very non-specific symptoms that are present in a wide range of benign conditions. Inevitably a large proportion of patients without cancer are investigated using this pathway. A number of issues have been raised by specialists in secondary care and GPs regarding the utility of this system, as the majority of cancer is still diagnosed via other routes. This includes non-TWW GP referrals and emergency presentations. The latest figures from 2016 show only 31% of CRC is diagnosed via the lower GI TWW pathway (282). Additionally, the system is extremely resource heavy, diverting time for primary care clinicians, specialists, endoscopists and administration staff away from other duties in order to meet operational standards, or face a financial penalty for the trust.

Despite the dominance of TWW work in both primary and secondary care, evidence for several key outcomes for this pathway are not available: for example cancer conversion rate and CRC stage at diagnosis are not collected nationally (148). NHS (National Health Service) England collects huge amounts of TWW data but this is not stratified into each cancer group; so the effect estimates calculated using this data may be unreliable. This information is vital to understanding the functionality of this pathway and whether it is meeting expectation.

In 2006, Thorne et al reviewed and evaluated the effect of the TWW on CRC and found low CRC detection rates, with no improvement in stage at diagnosis

compared to other referral routes (excluding emergency admissions) (283). The poor reporting quality and limitations in methodology of this review made it hard to replicate the study. For example, the report lacked a detailed search strategy, study inclusion criteria, study selection and extraction processes, quality assessment, and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) study flow diagram(284) .

Given the gaps in evidence discussed above, there is a need for an updated systematic review to comprehensively evaluate and collate all available published data and research evidence assessing the effect of a TWW system on the detection rate, stage at diagnosis, and mortality from CRC in England and Wales.

3.2 Objectives

The aim of this systematic review is to evaluate comprehensively the latest evidence on several key outcomes related to the two-week wait referral system in relation to suspected lower GI malignancy. The outcomes of interest for this review are as follows:

3.2.1 Primary outcomes

- The proportion of patients investigated for suspected CRC via the TWW system who are later diagnosed with CRC (cancer conversion rate).
- CRC stage at diagnosis.
- Effect of the TWW system on mortality from CRC.

3.2.2 Secondary outcomes

- Diagnoses made via TWW referral system other than CRC.
- Proportion of patients meeting the cancer waiting times standards for time to diagnosis and treatment i.e. a) proportion of patients seen within 14 days of referral, b) 31-day diagnosis to first treatment, c) 62-day wait for first treatment following GP urgent referral)

3.3 Method

This systematic review is reported according to recommendations from the Preferred Reporting Items for Systematic Review and Meta-analysis [PRISMA] 2009 statement[17]. Systematic review registration: PROSPERO CRD42016037368. The published protocol to this review can be located via PubMed ID::27784334

3.3.1 Study eligibility

All studies reporting the effects of the TWW system for suspected CRC on the detection rate, stage at diagnosis, and mortality from CRC were eligible for inclusion and those reporting on other diagnoses made and adherence to national TWW targets.

3.3.2 Study inclusion criteria

All UK-based reports addressing one or more of the primary or secondary outcomes of patients with suspected CRC referred through the TWW pathway, were included.

Study design: Prospective or retrospective cohort studies, case-control studies, cross sectional studies.

Study setting: Primary or secondary care based studies.

Population: Aged >18 years referred according to the NICE guidance for cancer referral from primary care for suspected CRC (note these guidelines have been amended since the TWW introduction therefore several versions were used between the studies depending on year of data collection).

Intervention: TWW referral system for suspected CRC. Studies reporting data on multiple speciality referrals via TWW were included if it was possible to separate outcome data by speciality to enable extraction of the CRC data only.

Comparator: Any comparator used that provides data on one or more of the primary or secondary outcomes. Studies with TWW data, but without a comparator arm were included.

Outcome: Rate of diagnosis, stage at diagnosis, and/or mortality from CRC. Studies that included data on secondary outcomes without primary outcome data were included.

Timing: Any length of follow up.

Language of publication: English only.

Date of publication: Eligible studies published in the year 2000 or onwards to correspond with implementation of the TWW referral system.

Type of publication: Full-text reports and abstracts.

3.3.3 Study exclusion criteria

Population: a) Patients with suspected cancer other than colorectal, b) anal cancer cases, c) participants aged <18 years, and d) pregnant women.

Intervention: The referral systems not related to the TWW pathway for CRC were excluded, for example urgent or routine GP referrals. Where it was not clear, attempts to contact the author were made. In cases where no response was gained the report was excluded as we could not be sure the results included only those from the TWW cohort. Non-UK based studies.

Outcomes: Studies not including data on any of the three primary outcomes (CRC diagnosis, stage, and/or mortality) or secondary outcomes (other diagnoses made, adherence to cancer pathway targets).

Study design: Reviews, editorials, letters, books, consensus statements or opinions. Review articles and letters were examined for identification of original studies but were not included if they contained no original data.

Date of publication: Reports published prior to the year 2000 were automatically excluded as the TWW pathway was not introduced until then.

3.3.4 Search strategy

Medline via Pubmed, Embase, TRIP and the British Library catalogue was searched from 1st January 2000 to 20th October 2017, using a combination of subject headings and keywords related to the TWW pathway and CRC. The search strategy is detailed in Appendix 1.

A series of search strategies was employed using increasingly refined search terms to increase the chances of identifying all relevant publications, as some focussed on TWW clinic outcomes for a range of specialities. Where multi-speciality studies occurred, the full publication was reviewed and if data could be extrapolated for CRC in isolation, it was included in the review. The search criteria were devised in collaboration with an information specialist at UHCW library and also peer reviewed by Professor R Arasaradnam and Professor N Waugh, both with systematic review experience.

Each reference list of eligible articles was hand searched to identify additional eligible reports/abstracts. Individualised searches on authors of included studies were employed. PROSPERO was also searched to ensure no other similar systematic review was being undertaken. Unpublished studies (grey literature) were sought by examining relevant conference proceedings. The database search was updated directly prior to the writing of this review to ensure that the latest evidence was included.

3.3.5 Study selection

All records were de-duplicated and compiled from a specialised database using Clinical Evidence Based Information Service (CEBIS) into an Excel file. Four independent reviewers were involved to ensure no delays at any stage of study selection due to other commitments or absences. Initially, all the identified titles/abstracts were screened by at least two independent reviewers (EM, MM, MT, MW) using a pre-designed and piloted form. Afterwards, the eligibility of all potentially relevant publications passing the abstract/title screening level was examined at full text level independently. Conflicts regarding study eligibility arising at both title/abstract and full-text screening levels were discussed and

resolved between the two reviewers or with the help of a third adjudicator. Further information was gathered from authors via email where questions regarding eligibility arose, if contact could not be made, the paper was excluded.

The study flow in terms of inclusion and exclusion is depicted in the PRISMA study flow diagram in figure 3.1.

3.3.6 Data extraction

Data extraction was performed using a pre-piloted data extraction sheet, and a minimum of three of a possible four reviewers completed data extraction on each eligible study independently (EM,MM,MW,MT). Disagreements between the reviewers were resolved via consensus. Where consensus could not be reached, additional opinion was sought from a senior reviewer (RA).

The data extracted included:

a) Study details (for example, author's name and year of publication, location, setting, number of centres, length of follow up and funding source).
b) Study participants (for example, recruitment dates, sample size, eligibility criteria, mean age, male/female ratio and ethnicity),
c) Study intervention.
d) Study outcomes (primary and secondary, and number of included patients and losses to follow-up). Where possible, missing statistical parameters of importance were calculated (for example, proportion, standard deviation standard error).

3.3.7 Quality assessment

Three independent reviewers used the National Institutes for Health study quality assessment tool for observational cohort and cross-sectional studies to appraise the methodological quality of studies (285). Assessment was based on 14 key questions with a yes/no answer, this allowed the reviewer to focus on the key concepts for evaluating the validity of the study, rather than a composite score or numerical value

that corresponded to an overall quality rating. The domains of bias covered by this tool include selection bias, information bias, measurement bias and also confounding. The original quality ratings of this tool (i.e. good, fair or poor), were re-categorised for this review as good, adequate or low quality and consensus between reviewers had to be reached. Any disagreements in the assessments between the reviewers were resolved by a consensus final decision from a senior reviewer (RA).

3.3.8 Data analysis and synthesis

The study and population characteristics were organised into summary tables and text. The synthesised data were organised and presented by primary and/or secondary outcome. For studies without a comparator intervention, the summary proportions for the post-referral cancer conversion, staging and diagnosis rates were calculated with corresponding 95% confidence intervals (95% CIs).

The summary measures were pooled across studies only if there was sufficient similarity in study design, settings, participants (age, sex, comorbidity), comparators (e.g. similar referral systems), length of follow-up and primary outcome measures (e.g. CRC diagnostic criteria, staging system and non-CRC diagnoses).

The pooled estimates of the proportions were generated using the random-effects meta-analysis. The heterogeneity and its extent for the pooled estimates was assessed via visual inspection of forest plots and statistical test results (I^2 statistic>50%).

For cancer stage, the Dukes staging system was used to allow comparison between groups, and where possible, TNM-reported stage was converted to Dukes. Forest plots were produced with collated staging data together with the national overall stage at presentation data to allow a comparator. Other diagnostic data was categorised and presented in a series of forest plots (e.g. polyp detection rate, non-CRC cancer diagnosis and other GI disorders).

Factors measured and/or reported heterogeneously were not pooled, but instead were synthesised narratively (e.g. adherence).

The meta-analyses were performed using an R package metafor (286). The raw proportions were transformed using Freeman-Tukey transformation during the estimation (287). Mortality data was incomplete in most cases therefore Kaplain-meier plots could not be produced and available results were narratively synthesised in a summary table.

The extent of publication bias was assessed by generating and inspecting funnel plot asymmetry as well as using the regression test for funnel plot asymmetry (288).

3.3.9 Deviations from the protocol

The review was conducted according to the previously published protocol (289). However, there have been minor deviations from the original methodology outlined in the protocol. Since devising the protocol the quality assessment strategy had to be altered as it became apparent that most included reports did not include a comparator group, therefore the quality assessment tool specified in the protocol was not appropriate as it was intended for use in studies with a comparator. An alternative validated quality assessment tool for observational cohort and cross-sectional studies was used. Low quality studies were included due to the observation early on in the review process, that most papers were deemed of low quality and exclusion would significantly reduce the number of included studies. This was felt to be detrimental to the quality of the review and the intended aim to summarise the available evidence.

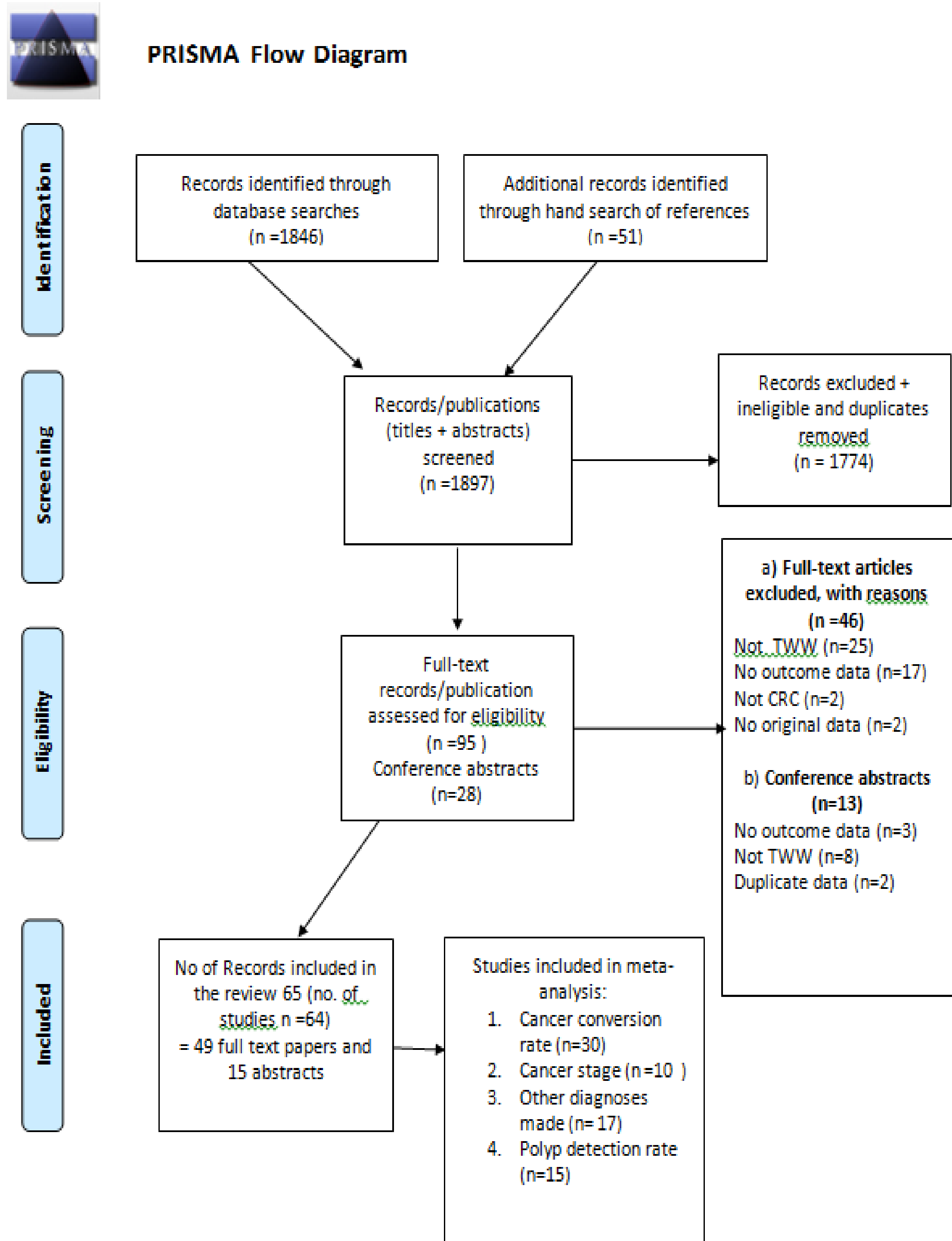
Many included studies were of low quality with a lack of demographic data and clear methodology. Conference abstracts had initially been excluded from this review, but it was judged that abstract data may also be of similar quality to that in full text papers and therefore complete exclusion was unjustified. As a result of this observation conference abstracts were included, but handled separately in the results section to avoid bias. Finally the intended data analysis was based on the assumption that most papers would provide a comparator group, as this was not the case, the data analysis was altered to account for this.

3.4 Results

3.4.1 Search results

The data collection period was 1st January 2000 to 31st October 2017. Between January 2016 and October 2017 a total of three separate searches (electronic databases, hand search and contacting authors) were undertaken and the results were collated. In total, 1897 bibliographic records were identified with duplicates and ineligible removed. Of these, 1655 were excluded as they were not conducted in the United Kingdom. 95 full paper records and 28 conference abstracts were reviewed in full. Of the 95 full text reports examined, 49 full-text papers met the inclusion criteria (references are listed in table 3.1) 15 of the 28 abstracts met inclusion criteria. Of the 49 included papers there are 50 sets of results as one publication reported two datasets from different centres. The total study population was 93655. The study selection process and reasons for exclusion during the full-text screening level are presented in fig 3.1 (the PRISMA Flow Diagram). A meta-analysis was performed on the studies reporting cancer conversion rate, cancer stage at diagnosis and other diagnosis made. Due to low volume and lack of consistency in presentation, meta-analysis was not conducted on patient survival outcome data or adherence to guidelines data. This has been summarised in narrative text.

Figure 3.1. PRISMA flow Diagram of study selection process



3.4.2 Studies and their characteristics

Of the 49 included studies, 34 were retrospective cohorts, 12 prospective cohort, one retrospective comparative, one audit and one review article which contained original data. Forty-seven studies were based in secondary care hospitals, two collected data both at primary and secondary care level (290, 291).

Two pairs of cohort studies (Bevis and Schneider plus Hall and Peacock) reported results in two different papers at differing time intervals (292-295). Original data on cancer conversion rate was followed up by mortality data several years later. Duplication was avoided by summarising the common data once, additional data was looked at separately as it represented new information on outcomes.

3.4.3 Population characteristics

Study population characteristics for the included studies are provided in Table 3.1. The studies have been split into those with and without a comparator group. Study details for each included conference abstract are found in Appendix 3.

In the majority of studies the setting was secondary care outpatients. The geographical catchment area related to that of the hospital where the study was performed and study population size varied accordingly from 42 (296) to 16431 (297) with total study population of 93,655. Only ten studies offered length of follow up data (293, 294, 298-305), which ranged from 6 months to 5 years.

In general, studies adopted one of two alternatives to data collection: 1) identify a set number or time-frame and look at all the CRCs diagnosed within these margins and track back to assess the referral mode to identify the TWW group. As this group looked at all cancers, they also tended to provide data on comparator groups too i.e. patients with CRC referred via non-TWW means. 2) identify a set number or time frame and look at all lower GI TWW referrals within these margins. Data sources utilised included administrative databases kept locally for TWW patients and local hospital cancer registries. Hospital notes were interrogated where required. The two studies that included primary and secondary care patients (ref) analysed cancer registries. Where it was not clear if the TWW group were all purely TWW, for

example when the term ‘fast-track’ was used, the reviewers sought clarification from the author. If there was no reply the study was excluded.

*Table 3.1: **Study and population characteristics:** With and without comparator group*

Study ID, year	Study characteristics	Population characteristics	Intervention groups	Outcomes	Methodological quality
		Studies without comparator			
Aljarabah, 2009 (306)	**Design:** Prospective, cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow-up:** n/s	**% Women:** n/s **Median age (range):** 75 (24-94) **Recruitment dates**: April 2006-Sept 2006 **Total sample size (baseline):** 217 **Total drop-out/excluded: n/s**	**Intervention group (n=217):** All lower GI tww referrals within recruitment period	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	low
Allgar, 2006 (307)	**Design**: Retrospective, cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 62 **Median age (range):** 66.7 **Recruitment dates**: 1st Jan 2001-31st Dec 2002 **Total sample size (baseline):** 764 **Total drop-out/excluded: 79**	**Intervention group (n=444):** All lower GI tww referrals within recruitment period, plus looked at all CRC's diagnosed within recruitment period.	1. Cancer conversion rate	low
Banerjea, 2017 (308)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** median 68.8 iqr 59-78.72 **Recruitment dates**: 1st aug 2014 - 30th Nov 2014 **Total sample size (baseline):** 553	**Intervention group (n=548):** All lower GI tww referrals within recruitment period	1. Cancer conversion rate 2. Other diagnosis made 3. Adherence to targets	low

		Total drop-out/excluded: 5			
Barwick, 2004 (309)	**Design**: Retrospective, cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 57 **Median age (range):** 68 (27-100) **Recruitment dates:** Jan 2001-Aug 2001 **Total sample size (baseline):** 149 **Total drop-out/excluded: 5**	**Intervention group (n=149):** All lower GI TWW referrals within recruitment period	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Acceptable
Beggs, 2011 (310)	**Design**: Retrospective, **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 36 **Median age (range):** 66 (30-100) **Recruitment dates:** Oct 2004- Sept 2005 **Total sample size (baseline):** 317 **Total drop-out/excluded: 0**	**Intervention group (n=317)**: All lower GI TWW referrals within recruitment period	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Acceptable
Bhangu, 2010 (298)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 514 days median (iqr 160-788)	**% Women:** 27 **Median age (range):** 71.6 (57-87.8) **Recruitment dates:** Jan 2006- July 2009 **Total sample size (baseline):** 1725 Total drop-out/excluded: 0	**Intervention group (n=1725, 108 with CRC analysed)**: All consecutive new CRC's diagnosed in TWW clinic within recruitment period	1. Cancer conversion rate	Acceptable
Borowski, 2016 (291)	**Design**: Retrospective **Study setting:** Somerset cancer registry (primary and secondary care) **No. of centres:** 44 **Length of follow up:** n/s	**% Women:** 41.3 **Median age (range):** 72 (22-94) **Recruitment dates:** Jan 2009-Sept 2014 **Total sample size (baseline):** 8273 **Total drop-out/excluded: 0**	**Intervention group (n=8273):** All CRC adenocarcinoma cases within cancer registry from GP practices that referred >10 CRC cases in total.	1. Cancer conversion rate	Low
Couch, 2014 (311)	**Design**: Retrospective	**% Women:** n/s	**Intervention group (n=**968**):** All consecutive	1. Cancer conversion rate	Low

	Study setting: Somerset cancer registry (primary and secondary care) **No. of centres: 1** **Length of follow up:** n/s	**Median age (range):** n/s **Recruitment dates:** All of 2013 **Total sample size (baseline):** 2950 **Total drop-out/excluded:** 0	CRC diagnosed via tww pathway in 2013 who had a colonoscopy as first line investigation		
Debnath, 2002 (312)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 2 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Aug 2000-July 2001 **Total sample size (baseline):** 239 **Total drop-out/excluded: 2**	**Intervention group (n= 239)**: All lower GI TWW referrals within recruitment period.	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Adherence to targets	Low
Eccersley, 2003 (313)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** A one year period (dates not included) **Total sample size (baseline):** 180 **Total drop-out/excluded: 7**	**Intervention group (n=173)**: All lower GI TWW referrals within recruitment period.	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Low
Gaunt, 2011 (314)	**Design**: Retrospective cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 62.2 **Median age (range):** 84 (80-96) **Recruitment dates:** Feb- July 2009 **Total sample size:** 777 (of which 159 >80 yrs) Total drop-out/excluded: 22	**Intervention group (n= 137)**: All lower GI TWW referrals within recruitment period >80 years old	1. Cancer conversion rate	Low
Glancy, 2004 (299)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 6 months after referral	**% Women:** n/s **Median age (range):** 70.8 **Recruitment dates:** Aug 2000-Nov 2001 **Total sample size:** 326	**Intervention group (n= 326)**: All lower GI TWW referrals within recruitment period except those with palpable rectal mass	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Low

		Total drop-out/excluded: n/s			
Hall, 2015 (294)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 8 months	**% Women:** 53 **Median age (range):** 69 (61-79) **Recruitment dates:** Nov 2011- Jan 2012, Feb-Apr 2012, May-July 2013 **Total sample size:** 1403 **Total drop-out/excluded: 53**	**Intervention group (n= 1350):** All lower GI TWW referrals within the three recruitment periods (pre, post and long-term post 'be clear on cancer' campaign(*Shared data with Peacock, 2013	1.Cancer conversion rate* 2. Other diagnoses made*	Acceptable
Hammond, 2007 (296)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** 8th Jan 2007-16th Feb 2007 **Total sample size:** 42 **Total drop-out/excluded: 0**	**Intervention group (n=42)**: All lower GI TWW referrals within the recruitment period if <80 years, no requirement for advocate and able to take bowel preparation	1. Cancer conversion rate 2. Other diagnoses made	Low
Khan, 2007 (315)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 59 **Median age (range):** 74 (62-81) **Recruitment dates:** Aug 2000-Dec 2003 **Total sample size:** 225 **Total drop-out/excluded: 0**	**Intervention group (n=225)**: All consecutive lower GI TWW referrals within the recruitment period who underwent CT Colonography.	1. Cancer conversion rate 2. Other diagnoses made	Acceptable
Khong, 2015 (300)	**Design**: Retrospective cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 2 years	**% Women:** 54 **Median age (range):** 67.5 (20-98) **Recruitment dates:** 1st Feb 2012- 30th April 2012 and 1st Sept 2012-31st Oct 2012 plus 'similar dates' in 2011 (Feb/Sept) **Total sample size:** 2349 **Total drop-out/excluded: n/s**	**Intervention group (n=2349)**: All consecutive lower GI TWW referrals within the recruitment periods.	1. Cancer conversion rate	Acceptable
Leung, 2010 (316)	**Design**: Retrospective **Study setting:** Secondary care	**% Women:** n/s **Median age (range):** n/s	**Intervention group (n=1100)**: All consecutive lower GI TWW referrals	1. Cancer conversion rate 2. Cancer stage at diagnosis	Low

	No. of centres: 1 **Length of follow up:** n/s	**Recruitment dates:** 12 month period (dates not stated) **Total sample size:** 1100 **Total drop-out/excluded: n/s**	within the recruitment period.	3. Adherence to targets	
Maruthachalam, 2005 (317)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 3 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Jan 2003- Dec 2003 **Total sample size:** 639 Total drop-out/excluded: 9	**Intervention group (n= 630)**: All consecutive lower GI TWW referrals within the recruitment period.	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Low
Mukherjee, 2010 (318)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres: 1** **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Feb 2007- Feb 2009 **Total sample size:** 662 **Total drop-out/excluded: 0**	**Intervention group (n= 662)**: All consecutive lower GI TWW referrals within the recruitment period.	1. Cancer conversion rate	Low
Padwick, 2013 (319)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 59 **Median age (range):** 68.4 (21-98) **Recruitment dates:** 2010 **Total sample size:** 940 **Total drop-out/excluded: 68**	**Intervention group (n= 940, smaller subset of 573 used for some data)**: All lower GI tww referrals within the recruitment period.	1. Cancer conversion rate 2. Other diagnoses made	Low
Panagiotopoulou, 2013 (320)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 2 **Length of follow up:** n/s	**% Women:** 53.8 **Median age (range):** 72 (61.9-80.9) **Recruitment dates:** Nov 2008- Jun 2009 and Apr 2010-Mar 2011 **Total sample size:** 1251 **Total drop-out/excluded:** 333	**Intervention group (n= 918)**: All consecutive lower GI tww referrals referred for anaemia only, within the recruitment period.	1. Cancer conversion rate	Acceptable
Patel, 2014 (321)	**Design**: Retrospective	**% Women:** 53.1	**Intervention group (n= 720)**: All consecutive lower	1. Cancer conversion rate	Acceptable

	Study setting: Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**Median age (range):** 68.5 (21-101) **Recruitment dates:** 1st Jan 2008- 31st December 2008 **Total sample size:** 720 **Total drop-out/excluded:** n/s	GI tww referrals within the recruitment period.	2. Cancer stage at diagnosis 3. Other diagnoses made	
Patel, 2016 (301)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** Up until Sept 2014 (variable time depending on when recruited)	**% Women:** 54.3 **Median age (range):** 45 (igr 41-48) **Recruitment dates:** Dec 2008- May 2014 **Total sample size:** 197 **Total drop-out/excluded:** 0	**Intervention group (n=197)**: All consecutive lower GI tww referrals aged <50 referred within the recruitment period.	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Survival data 4. Adherence to targets	Acceptable
Peacock, 2013 (295)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 53 **Median age (range):** 69 (61-79) **Recruitment dates:** Nov 2011 -Jan 2012 and Feb 2012-Apr 2012 **Total sample size:** 933 **Total drop-out/excluded:** 46	**Intervention group (n=887)**: All lower GI tww referrals referred within the two recruitment periods. (*shared data with Hall, 2015)	1. Cancer conversion rate* 2. Cancer stage at diagnosis 3. Other diagnoses made*	Acceptable
Rai, 2007 (322)	**Design**: Review article **Study setting:** not stated **No. of centres: 2** **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Not stated **Total sample size:** 1222 **Total drop-out/excluded:** n/s	**Intervention group (n= 1222)**: Not clearly stated	1. Cancer conversion rate 2. Adherence to targets	Low
Rai, 2008 (323)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 9 **Length of follow up:** n/s	**% Women:** 52.8 **Median age (range):** 68 (21-95) **Recruitment dates:** Sep 2003-Aug 2004 **Total sample size:** 1422 **Total drop-out/excluded:** 0	**Intervention group (n=509)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate	Acceptable

Raje, 2005 (324)	**Design**: Rectrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Not clearly stated **Total sample size:** 77 **Total drop-out/excluded:** n/s	**Intervention group (n=not clearly stated)**: CRC's diagnosed within an undefined timeframe in 2003	1. Adherence to targets	Low
Royle, 2014 (305)	**Design**: Retrospective **Study setting: S**econdary care **No. of centres:** 1 **Length of follow up:** 58 months (12-58) median 35	**% Women:** 54 **Median age (range):** 68 (18-96) **Recruitment dates:** Nov 2005-Nov 2009 **Total sample size:** 1690 **Total drop-out/excluded:** 84	**Intervention group (n=1606)**: lower GI tww referrals with left-sided symptoms only within recruitment period.	1. Cancer conversion rate 2. Other diagnoses made	Acceptable
Saratzis, 2015 (302)	**Design**: Retrospective, cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 20 months (14-26)	**% Women:** 42 **Median age (range):** 62 (sd 5) **Recruitment dates:** 1st Jan 2012-31st Dec 2012 **Total sample size:** 955 **Total drop-out/excluded:** 0	**Intervention group (n= 955)**: lower GI tww referrals who had declined or had a negative FOB as part of BCSP (aged 60-75)	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Survival data 4. Other diagnoses made	Acceptable
Shaw, 2009 (325)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 54.5 **Median age (range):** 58 (18-98) **Recruitment dates:** Aug 2005-Aug 2007 **Total sample size:** 2159 **Total drop-out/excluded:** 0	**Intervention group (n=2159)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate	Low
Shaw, 2008(326)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 58.3 **Median age (range):** 76 (40-97) **Recruitment dates:** Sep 2005-Sep 2006 **Total sample size:** 204 **Total drop-out/excluded:** 0	**Intervention group (n=204)**: Lower GI tww referrals with anaemia only within recruitment period	1. Cancer conversion rate	Low

Smith, 2007 (327)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 57 **Median age (range):** 66 (17-96) **Recruitment dates:** Jan 2002-Dec 2004 **Total sample size:** 2748 **Total drop-out/excluded:** 0	**Intervention group (n= 2748)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Adherence to targets	Acceptable
Sorelli, 2014 (328)	**Design**: Retrospective audit **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 68 **Median age (range):** 70 (35-86) **Recruitment dates:** Mar 2009-May 2009 **Total sample size:** 59 **Total drop-out/excluded:** 9	**Intervention group (n=50)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate 2. Other diagnoses made 3. Adherence to targets	Low
Spencer, 2004 (329)	**Design**: Prospective (presented as a letter) **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Jul 2000-Dec 2000 **Total sample size:** 243 **Total drop-out/excluded:** n/s	**Intervention group (n= 243)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate	Low
Vaughan-Shaw, 2013 (330)	**Design**: Retrospective cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 57 **Median age (range):** 72 (19-102) **Recruitment dates:** Jul 2007-Jul 2011 **Total sample size:** 2735 **Total drop-out/excluded:** 4	**Intervention group (n=2731)**: All lower GI tww referrals within recruitment period	1. Cancer conversion rate 2. Other diagnoses made	Acceptable
Vulliamy, 2016 (331)	**Design**: Retrospective cohort **Study setting:** Secondary care **No. of centres: 2** **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** April 2009 - April 2015 **Total sample size:** n/s	**Intervention group (n=n/s)**: All lower GI tww referrals within recruitment period	1. Adherence to targets	Low

		Total drop-out/excluded: n/s			
		Studies with comparator			
Aslam, 2017 (297)	**Design**: Retrospective, cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women: n/s** **Median age (range): n/s** **Recruitment dates**: Jan 2005 - Dec 2012 **Total sample size (baseline):**16431 **Total drop-out/excluded:** n/s	**Intervention group (n=1139):** All CRC's diagnosed via tww within the time frame. Also data on total tww population but not focus of the study. **Comparator group (n=2479)**: All CRC's diagnosed via non-tww within the timeframe	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Adherence to targets 4. Survival data	Low
Bevis, 2006 (292)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 50.7 **Median age (range):** 72 (25-96) **Recruitment dates:** Oct 2002-Sept 2004 **Total sample size:** 216 **Total drop-out/excluded:** 23	**Intervention group (n=97)**:All consecutive CRC's diagnosed within recruitment period via tww route **Comparator group (n=96)**: multiple a) non fast track referrals b)emergency presentations	1. Cancer stage at diagnosis 2. Adherence to targets	Acceptable
Chohan, 2004 (332)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** July 2000-December 2001 **Total sample size:** 462 (comparator total size not stated) **Total drop-out/excluded:** n/s	**Intervention group (n= 462)**: All lower GI TWW referrals within recruitment period **Comparator group (n= not clear)**: multiple a)non tww referral b) surgical/medical outpatient referrals c) medical inpatient referrals	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Other diagnoses made 4. Adherence to targets	Low
Courtney, 2013 (333)	**Design**: Retrospective, comparative analysis	**% Women:** n/s	**Intervention group (n= 9393)**: All lower GI TWW	1. Cancer conversion rate	Low

	Study setting: Secondary care **No. of centres:** 3 **Length of follow up:** n/s	**Median age (range):** n/s **Recruitment dates:** July 2006- Dec 2010 **Total sample size:** 12715 (all ages) **Total drop-out/excluded:** 0	referrals within recruitment period aged 60-79 only **Comparator group (n= 3322)**: Bowel cancer screening referrals within same recruitment period	2. Cancer stage at diagnosis	
Currie, 2012 (303)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** Median 5.03 years (0.91-5.89)	**% Women:** 49 **Median age (range):** 68 (39-78) **Recruitment dates:** Jan 2000-Dec 2005 **Total sample size:** 125 **Total drop-out/excluded:** 0	**Intervention group (n=52)**: All cases of rectal cancer referred via any route within recruitment period **Comparator group (n= 73)**: Routine lower GI referrals	1. Survival data 2. Adherence to targets	Low
Dua, 2009 (334)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 70 **Median age (range):** 69 (24-88) **Recruitment dates:** not stated **Total sample size:** 150 **Total drop-out/excluded:** 0	**Intervention group (n= 75)**: Cases of CRC diagnosed via lower GI TWW clinic **Comparator group (n=75)**: Non-tww GI referrals that subsequently are diagosed with CRC	1. Cancer stage at diagnosis 2. Adherence to targets	Acceptable
Flashman, 2004 (335)	**Design**: Prospective audit **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 60 **Median age (range):** 70 (25-93) **Recruitment dates:** 1st July 2000- 30th June 2001 **Total sample size:** 2573 **Total drop-out/excluded:** 63	**Intervention group (n= 695)**: All lower GI TWW referrals within recruitment period. **Comparator group (n=1815)**: Routine lower GI surgical outpatient referrals	1. Cancer conversion rate 2. Cancer stage at diagnosis 3. Adherence to targets	Low
John, 2008 (290)	**Design**: Retrospective **Study setting:** Primary and Secondary care **No. of centres:** 1	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** Not stated	**Intervention group (n= 100)**: All consecutive lower GI TWW referrals within the recruitment period.	1. Cancer conversion rate 2. Other diagnoses made	Low

	Length of follow up: n/s	**Total sample size:** 200 **Total drop-out/excluded:** 0	**Comparator group (n= 100)** : Routine consecutive lower GI clinic referrals		
Neal, 2007 (304)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** Up until June 2003	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** 2000-2001 (total 2 years) **Total sample size:** 239 **Total drop-out/excluded:** 22	**Intervention group (n= 51)**: All CRC's diagnosed within the recruitment period for all referral routes. **Comparator group (n= not clear) :** All non-tww CRC cases within recruitment period.	1. Cancer stage at diagnosis 2. Survival data 3. Adherence to targets	Low
Rao, 2005 (336)	**Design**: Prospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** June 2003-December 2003 **Total sample size:** Not clear **Total drop-out/excluded:** n/s	**Intervention group (n= 319)**: All lower GI tww referrals within recruitment period **Comparator group (n= not clear):** Non-tww referral pathways	1. Cancer conversion rate 2. Adherence to targets	Low
Schneider, 2012 (293)	**Design**: Retrospective **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** 5 years	**% Women:** 50.7 **Median age (range):** 73 **Recruitment dates:** Oct 2002- Sep 2004 **Total sample size:** 193 **Total drop-out/excluded:** 4	**Intervention group (n= 96)**: All CRC patients referred via any route with recruitment period **Comparator group (n= 93):** a) Non-tww referral pathways b)emergency presentation	1. Survival data *Linked with Bevis, 2006	Acceptable
Thornton, 2016 (337)	**Design**: Retrospective cohort **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** 44.4 **Median age (range):** 73 **Recruitment dates:** Apr 2006-Dec 2012 **Total sample size:** 558 **Total drop-out/excluded:** 0	**Intervention group (n= 197):** All CRC's diagnosed within recruitment period under care of two designated consultants **Comparator group (n= 361):** a)non-tww routes of referral b)acute admissions	1. Survival data	Acceptable

Walsh, 2002 (338)	**Design**: Prospective audit **Study setting:** Secondary care **No. of centres:** 1 **Length of follow up:** n/s	**% Women:** n/s **Median age (range):** n/s **Recruitment dates:** 1st Aug 2000-31st Oct 2000 **Total sample size:** 229 **Total drop-out/excluded:** not clear	**Intervention group (n= 105)**: All lower GI referrals within recruitment period **Comparator group (n=124)**: a) non-tww referral pathway with lower GI symptoms	1. Other diagnoses made 2. Adherence to targets	Low

3.4.4 Studies including a subset population

Eleven of the 49 papers reported outcome data on only a subset of the two-week wait cohort. Table 3.2 shows the subset papers by author, together with the subgroup category.

***Table 3.2: Summary of all papers reporting only TWW subset data** and description of the subgroup.*

Author, year	Subgroup category	Number of full text papers
Courtney 2013	Aged 60-79	1
Currie, 2012	Rectal cancer cases only	1
Gaunt, 2011	Aged >80	1
Glancy, 2004	Only investigated using barium enema	1
Hammond, 2008	Aged <80	1
Khan, 2007	Only investigated using CT colonography	1
Panagiotopoulou,2013 **Shaw, 2008**	Referred for anaemia only	2
Patel, 2016	Aged <50	1
Royle, 2014	Left sided symptoms only	1
Saratzis, 2015	Aged 60-75	1

3.4.4.1 Age- restricted cohort

Six papers included patients within a defined age range. These were categorised as follows; a) aged <50 years b) Aged 60-75 years c) Aged 60-79 years d) aged <80 years e) aged >80 years. The studies by Saratzis, 2015 and Courtney, 2013,

collected TWW data on a pre-defined age group in order to age-match this group with the Bowel Cancer Screening Programme (BCSP) population.

3.4.4.2 Symptom-restricted and image-restricted cohorts

Two studies included only the patients referred using the anaemia criteria. Royle, 2014, excluded those with anaemia and palpable right sided mass. Two studies focussed on using one imaging modality only: Glancy, 2004, used a barium enema and Khan, 2007, investigated the use of CT colonography, with particular focus on extra-colonic pathology.

Through consensus discussion between reviewers, it was agreed to include these TWW subset papers in the review as they included important results. However as these studies were not representative of the TWW cohort as a whole, and are therefore liable to distort results, they were analysed separately. Therefore unless otherwise stated in this results section, the eleven papers presenting subset populate data are not included in the main analysis.

3.4.5 Quality of included studies

The methodological quality of studies are provided in Table 3.1. No study was deemed to be of high quality. Nineteen of the 49 studies were of acceptable and remaining 30 were of low quality. Poor reporting of study population characteristics (item 2) and poorly defined outcomes (item 11) were the main contributory factors leading to the judgement of low quality. In the studies with a comparator, it was not clear what differences there were between patient groups regarding sex and age, making it difficult to assess a potential bias due to confounding. Conference abstracts were all judged to be of low quality due to the lack of comprehensive methodology possible due to a restricted word count policy.

3.4.6 Synthesis of outcome data

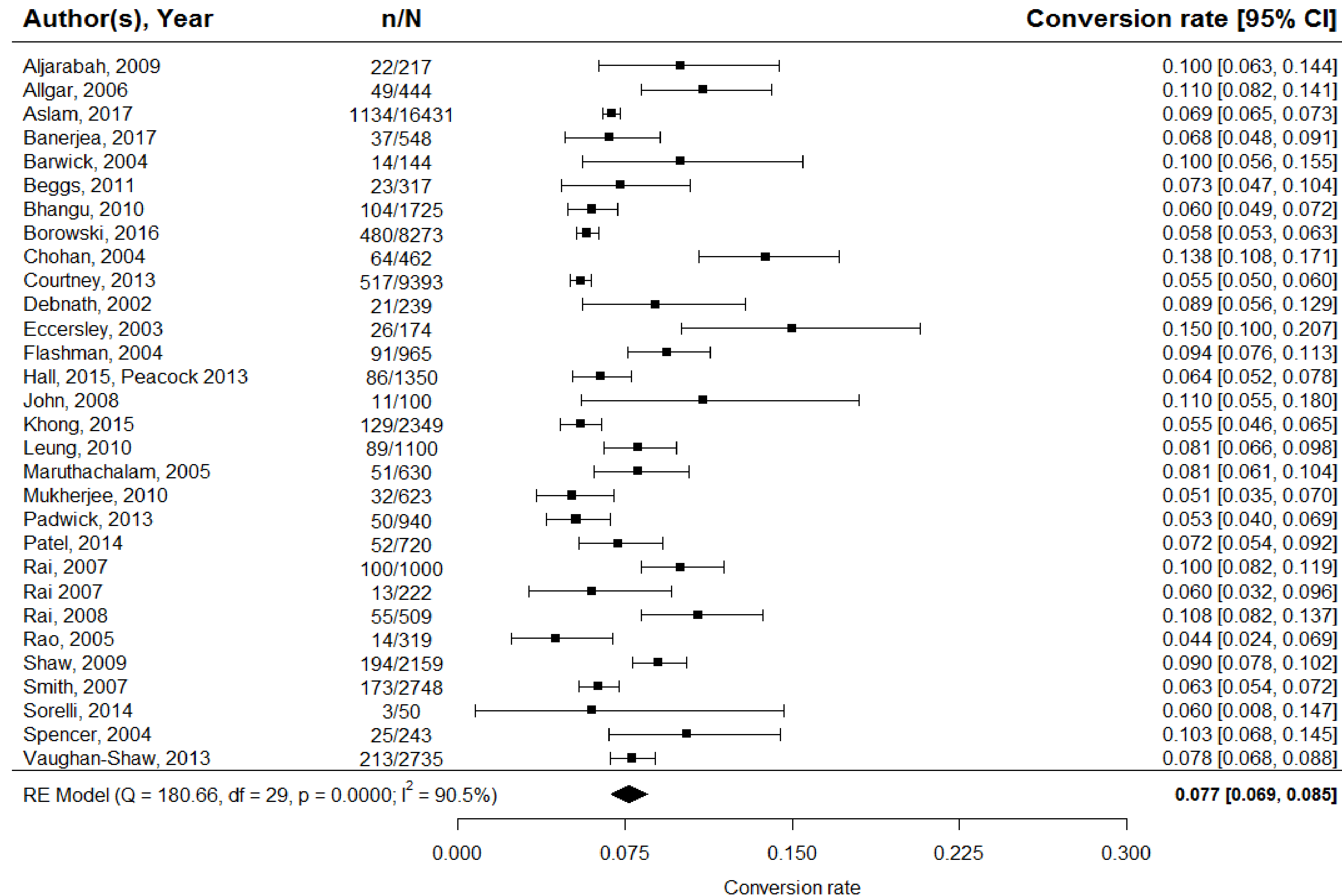
Results are reported according to outcome, as set out in the method.

3.4.7 Cancer conversion rate

A total of 38 (77.5%) full-text study reports provided cancer conversion rates with a total study population of 70569. Nine (18%) of these provided data on a subset of patients (see table 3.2). 13 (86%) conference abstracts reported cancer conversion rate. Rai et al reported two separate cancer conversion rates, resulting in 52 values from 51 texts.

The cancer conversion rate was reported as a total number, together with a percentage. Weighted mean cancer conversion rates were 8.4% (95% CI: 7.5-9.3) for the total group, 7.7% (95% CI: 6.9-8.5) for full papers, 8.2% (95% CI: 5.5-11.3) for the subsets and 10.4% (95% CI: 7.8-13.2) for abstracts. Figure 3.2 shows the results as a forest plot for the full paper reports with overall pooled cancer conversion rate using a random effects model. The forest plots for cancer conversion rate for the subset and abstract groups are shown in appendix 4.

***Figure 3.2. A forest plot of cancer conversion rate** for full paper reports with corresponding 95% CI. Overall pooled conversion rate using random effects model is shown at the bottom of the plot.*

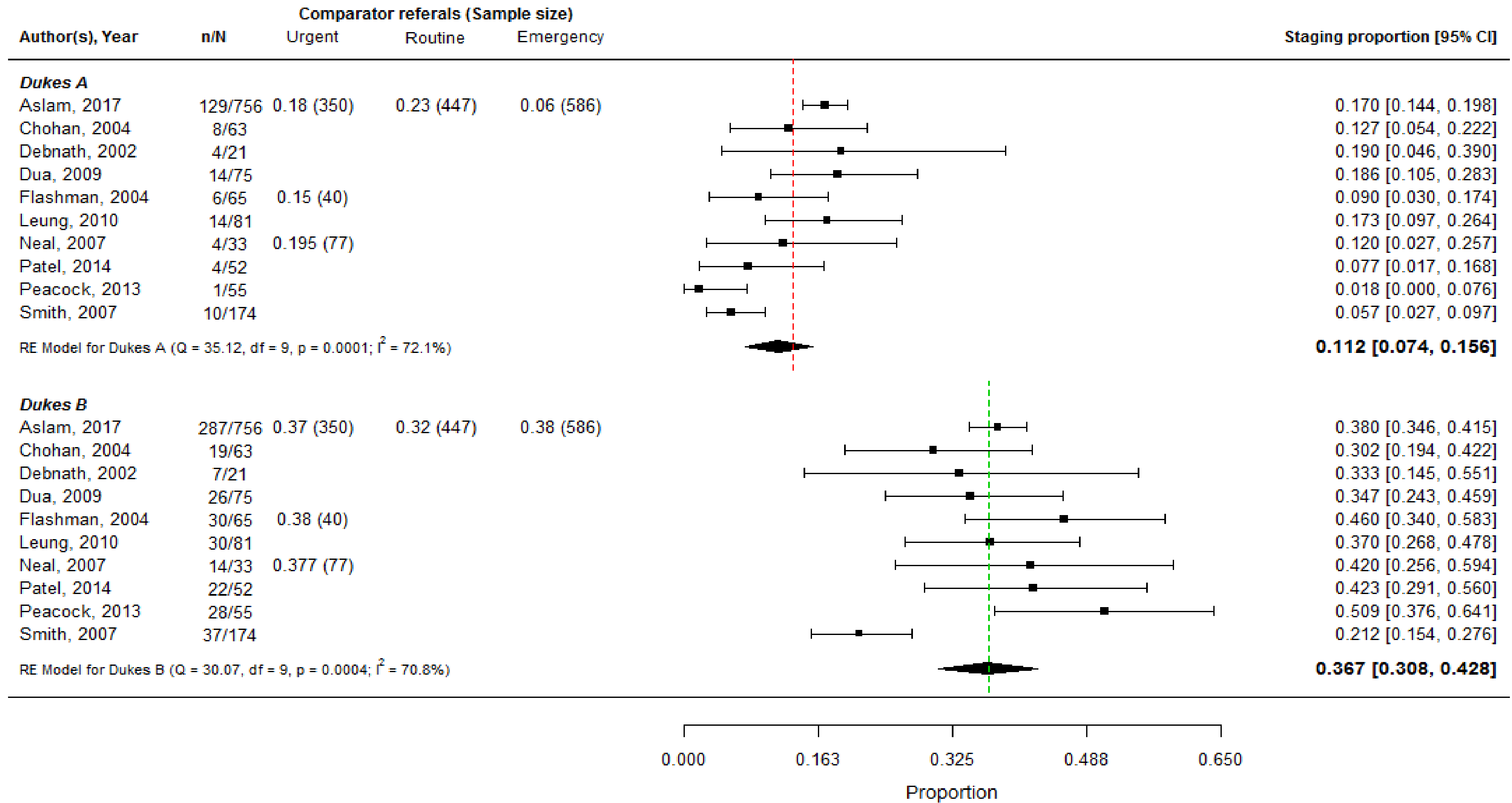


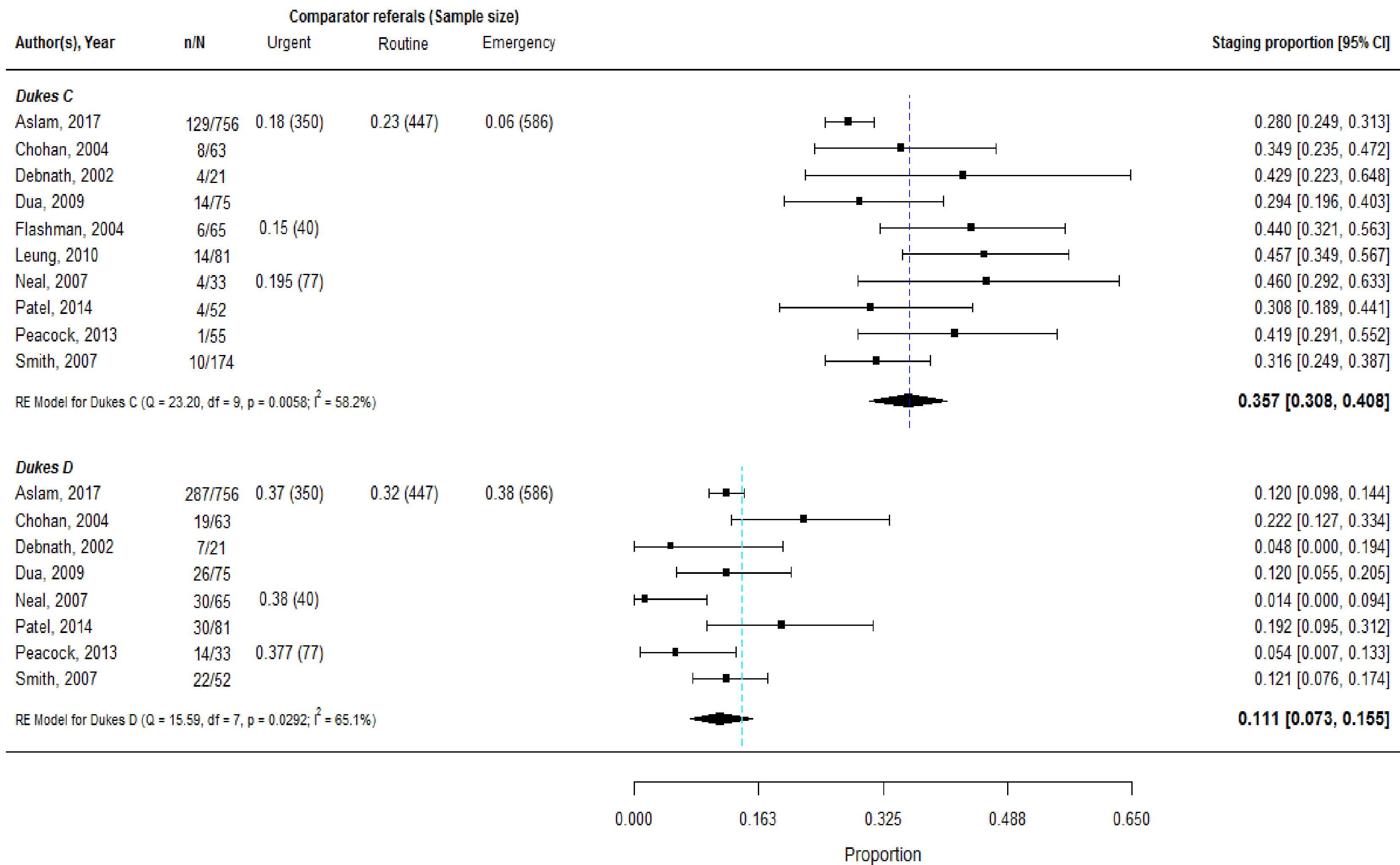
3.4.8 Cancer stage data from studies with no comparator group

The CRC stage was reported in 14 (28.5%) papers, 3 (21.4%) of which provided data only on a subset and therefore were analysed separately. Total study population was 39678. Of the reports that analysed an unselected TWW cohort, there were 9 reporting Dukes staging. In 1 case, it was possible to convert TNM staging to Dukes (339) equating to a total of 10 staging results. Dukes D data was not reported in two of the studies. Pooled data shows the proportion presenting at Dukes A = 11.2% (95% CI: 7.4-15.6), B=36.7% (95% CI: 30.8-42.8), C= 35.7% (95% CI: 30.8-40.8) and D=11.1% (95% CI: 7.3-15.5) (figures 3.3a and 3.3b). Data from the National Cancer Registration and Analysis Service (282) on Dukes staging at presentation for all CRC in England, are plotted using a vertical coloured line on each of the Dukes stage forest plots. This forms a comparison with national data on Dukes stage at presentation, corresponding values are: A=13.2%, B= 36.9%, C=35.9%, D= 14.0%. The vertical lines transecting the CI bars in each Dukes stage plot indicate that there is no statistically significant difference between the distribution of stage at diagnosis in the TWW cohort, compared with using any route combined.

Where the subset data was analysed, Courtney et al was the only study to report Dukes staging data and in the 60-79 age group only. Patel et al, who looked at the < 50 year age group only, reported that 22.2% had metastatic disease at presentation. In comparison Saratzis et al considered only patients aged 60-75 who had declined, or had a negative FOB, and found 19.2% of patients had metastatic disease at diagnosis.

Figure 3.3a and 3.3b Forest plot of papers reporting proportions presenting at Dukes stage A and B (figure 3.3a) and stages C and D (figure 3.4b). *The coloured vertical lines represent the proportions presenting at each stage via any non-emergency route. Where reported, comparator group data is stated numerically.*





Comparator referals (Sample size)
Author(s), Year
n/N
Urgent
Routine
Emergency
Staging proportion [95% CI]
Dukes C
Aslam, 2017 129/756 0.18 (350) 0.23 (447) 0.06 (586) 0.280 [0.249, 0.313]
Chohan, 2004 8/63 0.349 [0.235, 0.472]
Debnath, 2002 4/21 0.429 [0.223, 0.648]
Dua, 2009 14/75 0.294 [0.196, 0.403]
Flashman, 2004 6/65 0.15 (40) 0.440 [0.321, 0.563]
Leung, 2010 14/81 0.457 [0.349, 0.567]
Neal, 2007 4/33 0.195 (77) 0.460 [0.292, 0.633]
Patel, 2014 4/52 0.308 [0.189, 0.441]
Peacock, 2013 1/55 0.419 [0.291, 0.552]
Smith, 2007 10/174 0.316 [0.249, 0.387]
RE Model for Dukes C (Q = 23.20, df = 9, p = 0.0058; I² = 58.2%) 0.357 [0.308, 0.408]
Dukes D
Aslam, 2017 287/756 0.37 (350) 0.32 (447) 0.38 (586) 0.120 [0.098, 0.144]
Chohan, 2004 19/63 0.222 [0.127, 0.334]
Debnath, 2002 7/21 0.048 [0.000, 0.194]
Dua, 2009 26/75 0.120 [0.055, 0.205]
Neal, 2007 30/65 0.38 (40) 0.014 [0.000, 0.094]
Patel, 2014 30/81 0.192 [0.095, 0.312]
Peacock, 2013 14/33 0.377 (77) 0.054 [0.007, 0.133]
Smith, 2007 22/52 0.121 [0.076, 0.174]
RE Model for Dukes D (Q = 15.59, df = 7, p = 0.0292; I² = 65.1%) 0.111 [0.073, 0.155]
0.000
0.163
0.325
0.488
0.650
Proportion

3.4.9 Cancer stage from studies with a comparator (non-TWW) group

Five groups also reported staging data on non-TWW cohorts for comparison. The proportion of patients presenting at each stage by comparator group is listed within figures 3.3a+b but are also shown in table 3.3. On the forest plot, numbers in brackets represent sample size. The comparator groups were not consistent between studies. Aslam, Flashman and Neal analysed Dukes stage on an unselected TWW cohort, Courtney et al only looked at a subgroup aged 60-79 and provided comparative staging data on the Bowel Cancer Screening Programme patients.

Table 3.3. Staging data from comparator (non-TWW) groups. *Any statistical analysis provided by the author is also stated.*

Study ID, year	Dukes staging		Statistical analysis
Aslam, 2017	**TWW**	**Comparator (urgent referrals)**	Calculated OR's for likelihood of having metastatic disease
	A: 128 (17%)	A: 63 (18%)	tww vs Routine OR:1.33 CI: 0.98-1.54 p=0.072
	B: 287 (38%)	B: 127 (37%)	
	C: 213 (28%)	C: 108 (31%)	
	D: 90 (12%)	D: n/s	BCSP vs tww OR: 0.34 CI: 0.27-0.47 p=<0.001
	u/k: 38 (5%)	u/k 14 (4%)	
Flashman, 2004	**TWW**	**Comparator (non-tww outpatient clinic)**	None performed
	A: 9%	A: 15%	
	B: 46%	B: 38%	
	C: 44%	C: 46%	
	D: n/s	D: n/s	
Dua, 2009	**TWW (converted)**	**Comparator**	Comparison between proportion with stage i/ii vs stage iii/iv in TWW and non-TWW was not significantly different p=0.87
	A: 14 (18.6%)	A: 13 (17.4%)	
	B: 26 (34.7%)	B: 30 (40%)	
	C: 22 (29.4%)	C: 18 (24%)	

	D: 9 (12%) u/k: 4 (5.3%)	D: 14 (18.6%) u/k: 0	
Neal, 2007	**TWW** A: 4 (12%) B: 14 (42%) C: 15 (46%) D: 0	**Comparator (Non-tww lower GI referrals)** A: 15 (19.5%) B: 29 (37.7%) C: 33 (42.9%) D: 0	No significant difference in TWW vs non-TWW Dukes staging p=0.683
Courtney, 2013 (Aged 60-79 years only)	**TWW** A: 45 (15.4%) B: 92 (31.5%) C: 86 (29.5%) D: 62 (21.2%)	**Comparator-BCSP** A: 180 (50.6%) B: 79 (22.2%) C: 77 (21.2%) D: 16 (4.5%)	Proportion of tumours at each Dukes stage were significantly different between TWW and BCSP groups i.e. significantly more Dukes A in BCSP group and more Dukes D in TWW group p=< 0.001

The comparator data demonstrated there was more early stage disease in the BCSP compared with the TWW group with a significantly reduced OR. There were also fewer patients presenting with Dukes A in the emergency presentation group (8% vs national average of 13.3%).

3.4.10 Other information on cancer stage

Bevis et al, reported TNM staging data on those who underwent surgery with curative intent only, but it was not in a format that could be converted to Dukes staging. They compared TWW data with those presenting to outpatients via non-TWW modes and emergency presentations. They did not find a statistically significant different in the proportion presenting with metastatic disease between the three groups (p=0.86).

3.4.11 Other diagnosis

This includes polyp detection rate, IBD, diverticular disease, non-colorectal cancer and other diagnoses (normal, IBS or haemorrhoids)

3.4.12 Polyp detection rate

15 (30.6%) papers reported on polyp detection rate with a combined study population size of 7936. Only 6 stated that polyps detected were adenoma type, the remaining 9 did not state polyp type and therefore have been labelled as 'undefined'. Pooled weighted polyp detection rates were calculated both for adenomas and polyps undefined separately and then as one group. Proportions were compared with the adenoma detection rate (ADR) found in the BCSP population (the figure was taken from the Coventry and Warwickshire regional bowel screening data for 2017); they were also compared with the UK global ADR (27, 340) (figure 3.4).

The report by Khan et al was excluded from combined analysis as they only performed CTC on patients therefore the results are not representative of the standard TWW investigation pathway and have the potential to introduce unnecessary bias.

ADR was 8.5% (95% CI: 6.6-10.6), polyp undefined detection rate was 12.6% (95% CI: 9.3-16.3) and the detection rate for all polyps combined was 10.8% (95% CI: 8.5-16.3). All are lower than the ADR from colonoscopies via all routes combined and via the BCSP of 15.9% and 50.8% respectively.

3.4.13 Non-CRC malignancy and other GI diagnoses

17 papers reported on other diagnoses with a total sample size of 8150 (figure 3.5). Diagnoses were separated into non-CRC cancer, IBD, diverticular disease and other (normal, IBS or haemorrhoids). Diagnosis of non-CRC cancer and IBD were very low at 3.2% (95% CI :0.9-6.6) and 3.9% (95% CI: 2.9-4.9) respectively. Overall, diverticular disease was seen in 23.4% (95% CI: 18.9-28.1) of cases and 54. 6% (95% CI: 46.2-62.8) fell into the other category.

Figure 3.4. Forest plot of adenoma and undefined polyp detection rate *with 95% CI. Red vertical line represents national ADR for all procedures, green vertical line represents ADR for the BCSP cohort.*

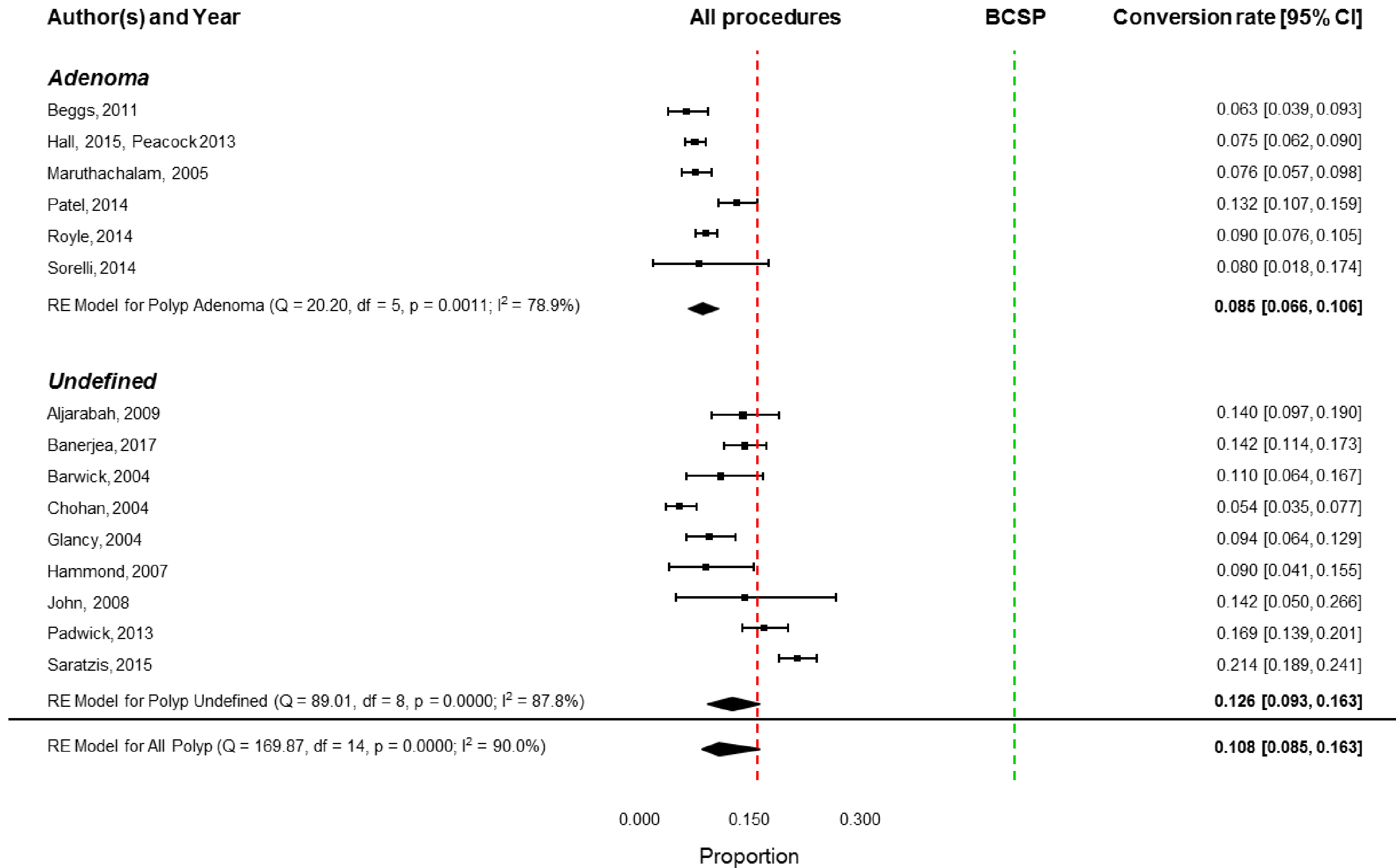
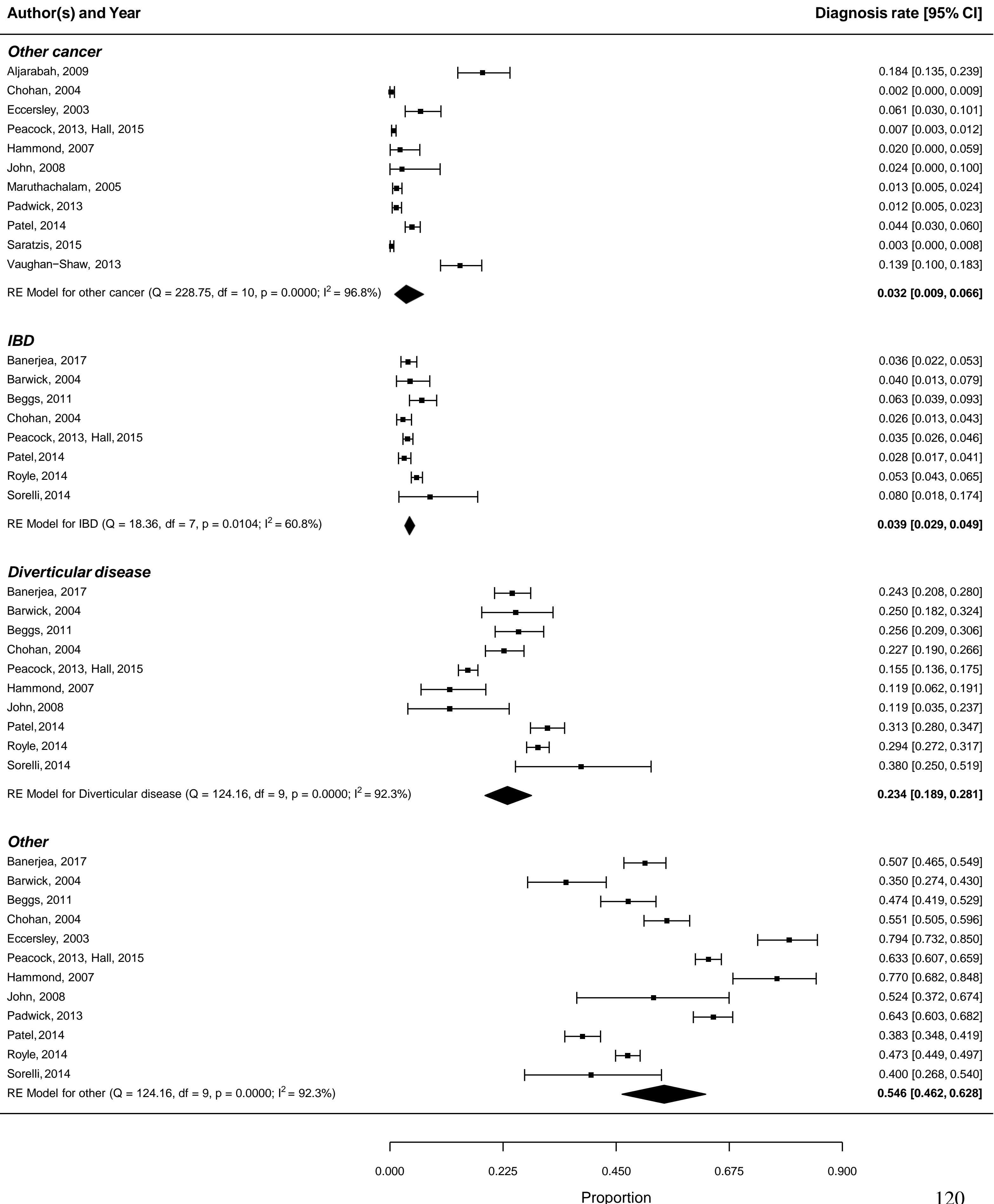


***Figure 3.5. Forest plot of non-CRC diagnosis rates** with 95% CI. This includes a) other cancer (non-CRC), b)IBD, C) diverticular disease and d) Other (normal, IBS or haemorrhoids)*



3.4.14 Adherence

Of those papers reporting on adherence to the TWW targets, the total study population was 9227. Data was categorised based on the operational standards set out by NHS England. Data reporting on compliance to standards is summarised in table 3.4. No statistical analysis was possible due to the small volume of data and heterogeneity between result reporting.

Table 3.4. Summary of all data reported on adherence to operational standards for TWW referrals. *Alphabetised by first author, time range in brackets where stated.*

	TWW			Routine		
Author, year	**Median days to specialist clinic**	**Median days to diagnosis**	**Median days to treatment**	**Median days to specialist clinic**	**Median days to diagnosis**	**Median days to treatment**
Aslam, 2017	12	13	63	35	55	112
Banerjea, 2017		24.5				
Barwick, 2004	10 (0-54)					
Bevis, 2006	31 (5-127)		42.5 (5-173)	69 (9-1021)		57.5 (7-273)
Chohan, 2004	12 (2-28)		36 (9-134)	24 (1-118)		36.5 (1-226)
Currie, 2012		57 (31-101)			74 (31-122)	
Dua, 2009	10		74	57		161
Eccersley, 2003	10 (1-47)					
Flashman, 2004	12			27		
Patel, 2016	11 (iqr 7-13)					
Rao, 2005	13 (0-276)			25		
Vulliamy, 2016	11 (8-14)		58 (49-62)			
Walsh, 2002	8.64	47				
			Abstract			
Gandy, 2002	9.7 (sd 5.1)					

In general, the median days to specialist clinic target met the 93% operational standard in a high proportion of cases. Conversely, days from referral to treatment initiation were associated with median days outside of the desired timeframe (exact proportions cannot be calculated from the available data).

3.4.15 Survival

Only 7 papers reported any survival data with few consistent methods of data reporting, therefore meta-analysis was not possible. The total study population was 18698. Table 3.5 provides a summary of the available information.

Table 3.5. Studies that reported survival rates in the TWW vs non-TWW cohorts.

Author, year	TWW		non-TWW	
	2-year survival	5-year survival	2 year survival	5 year survival
Schneider, 2012		60.3% +/- 5.2		58.8 +/-5.3
Currie, 2012		49% (1521 days)		52% (1591 days)
Thornton, 2016	74.10%	52.30%	73.9% (tww vs non-tww p=0.837)	53.8% p=0.889 (tww vs non-tww)

2-year survival was only reported by Thornton, with 74.1% in the TWW cohort and 73.9% in the non-TWW group alive at two years, p=0.837. They also found no statistically significant difference in 5-year survival between the TWW and non-TWW groups, with survival at 52.3% and 53.8% respectively, p=0.889. Mean survival in this group was 52.7 months in the TWW cohort and 64.1 months in the non-TWW group p=0.912. Unsurprisingly the patients presenting acutely in this study had poorer survival trends with 2-year survival of 38.2% and 5-year survival of 23.7%.

5-year survival was reported by Schneider as 60.3% in the TWW group and 58.8% in the non-TWW group. Mean survival was also similar; 43.5 months for TWW

and 43.7 months for non-TWW. In this study patients who presented acutely had 43.5% chance of 5-year survival with mean survival 33.7 months. 5-year survival in the study by Currie was 49% in the TWW group and 53% in the non-TWW group

Aslam found a median survival of 3.5 years in the TWW group, 5.4 years in those diagnosed via the routine pathway, 3 years in those referred urgently and 1.1 years in those presenting as emergencies.

Few conclusions can be drawn based on this small amount of survival data.

3.4.16 Assessment of heterogeneity and publication bias

Heterogeneity was generally high for all reported outcomes, as calculated by I^2. The main reason for heterogeneity was considered to be related to sample size, as stated previously, sample size was highly variable. However, this theory was tested by splitting the cancer conversion outcome data according to study sample size. Medium sample size was 630 patients, there were 16 papers below this sample size and 14 above. I^2 values were 72% and 90% respectively. The two corresponding forest plots are contained in appendix 5. Therefore only a modest reduction in heterogeneity was demonstrated. In this case, other causes for heterogeneity are unknown. Although demographic data reporting was inconsistent, where stated, the results for study population age and sex ratios were similar. In addition, the methodology behind the study design was similar. A funnel plot was not possible as there were too few studies comparing the two-week wait pathway with a comparator group

3.5 Discussion

In this systematic review and meta-analysis the available data on key outcome measures for the lower GI TWW referral pathway has been summarised and analysed in 49 studies and 15 abstracts with a total study population of 93,655.

3.5.1 Cancer conversion rate and stage at diagnosis

3.5.1.1 Major findings

The pooled cancer conversion rate was 7.7%. This is comparable to the cancer detection rate in the asymptomatic BCSP population of 8% (2017 data) and applies to a similarly aged cohort: median age of the patients in this study was 72 years, compared with the 60-75 age group for cancer screening. This result is also consistent with the overall cancer conversion rate for all cancers diagnosed through any TWW pathway which is 7.6%(148). Although a large amount of data is collected on a quarterly basis on TWW pathways by NHS England, as mentioned previously, the national TWW performance data for cancer conversion rate is not kept for each individual cancer. Data can be extrapolated to give an estimate of the conversion rate for each site, but there is the likelihood of introducing error as cancer diagnosis is calculated from the total number meeting the target to start treatment within 62 days. This is likely to be inaccurate as there is no definite way of ensuring cases that don't meet the target, but have cancer, or those not treated are not missed. In addition, it is not entirely clear patients included in the figures have a diagnosis of colon adenocarcinoma as it is not possible to cross check individual cases with histology. Due to the above gaps in accurate data collection, results of this review provide the largest and most comprehensive dataset on cancer conversion rate in the lower GI TWW population.

This is the first report to document staging data for the TWW cohort. Overall stage at diagnosis for Dukes A-D was 11.2%, 36.7%, 35.7% and 11.1% respectively. When compared with the national staging data (341) no statistically significant difference in stage distribution was seen in the TWW cohort compared with CRC diagnosed via any route combined. Nearly half of patients present with stage three or four disease with corresponding survival of 47.7% and 6.6% respectively (342).

According to the National Cancer Intelligence Network (NCIN) report on colorectal staging data quality (343), nationally collected data is poorly reported and incomplete. Stage at diagnosis is directly linked to survival, results from this review indicate that the TWW process is not having a significant impact on improving CRC survival.

3.5.1.2 Strengths and limitations of the evidence

This review encompasses a large unselected population size with a high number of reports evaluated from a wide range of centres nationally, this should theoretically, provide data on a far reaching socio-economic and ethnically diverse population, although this cannot be confirmed as this data was not presented by any study. As discussed previously, the general quality of papers was predominantly poor. Cancer conversion rate is a hard endpoint: it is therefore reasonable to assume values were accurate and comparable unless, where stated, a subset was used. The value of abstract data was limited as it appeared to be subject to significant bias resulting in a higher cancer conversion rate than seen from the full paper reports. Several reasons could explain this. Firstly, most of this abstract data was produced in the one to two years following TWW implementation, therefore it is possible that the referral criteria were more rigorously adhered to hence the higher cancer detection. Also there is likely to be a degree of lead time bias due to the implementation of a new system for cancer detection, so a transiently higher incidence would be expected. It has been acknowledged in this review that heterogeneity is high between the studies, and cannot be fully explained by sample size. This must be taken into consideration when assessing the results.

3.5.1.3 Consistency of the findings

There has been only one other review looking at TWW outcomes in lower GI cases. As highlighted in the introduction Thorne et al left some gaps in methodology and evidence collection. They found a cancer detection rate of 10.3%. The higher detection rate may relate to the year of publication of included studies. It was performed close to the initiation of the TWW pathway, with the same bias that applied to the conference abstract results playing a role. Additionally, the referral

criteria have since been widened. This has resulted in an increase in the number of patients eligible for referral and a subsequent increase in the number of non-CRC cases being investigated.

3.5.2 Other non-CRC diagnoses made and adherence to referral targets

3.5.2.1 Major findings

There were a high proportion of non-CRC diagnoses made through the TWW pathway, which reiterates the fact that lower GI symptoms are largely non-specific and present in a variety of GI complaints. ADR was 8.5%, much lower than the BCSP rate at 50.8% or the overall ADR via all routes to colonoscopy at 15.9%. The lower ADR may be related to year of data collection, as many are over ten years old. There has been a concerted drive in recent years through targets and guidelines to use ADR as a marker of endoscopist performance (340). Endoscopists receive their individual ADR and this is monitored by the Joint Advisory Group (JAG) against expected targets. Advances in scope technology and bowel preparation have added further to improvements in polyp detection.

The presence of adenomas probably represent incidental findings rather than being related to symptomatology in the majority of cases, most small polyps are unlikely to cause symptoms.

Other diagnoses were split into other non-CRC cancers, IBD, diverticular disease and other (comprised of IBS, normal and haemorrhoids). Cases of IBD and other cancers were low and over half the patients either had a diagnosis of normal, IBS or haemorrhoids. This reiterates the poor predictive value of the symptomatic patient in the diagnosis of significant lower GI pathology. NHS England makes no comment on the probability or expectation that non-CRC diagnoses will be made through the TWW pathway, but these results indicate that a patient is more likely to have no pathology than any other diagnosis combined.

Adherence data was heterogenous and comparisons were difficult as the median days to target does not make it clear if operational standards were met. There were widely varying ranges within each result with very little adherence data on target to

starting treatment. It is possible that this data is harder to capture from computer databases. In summary, no firm conclusions on adherence data can be made from this review due to the paucity of data and variations in reporting technique.

3.5.2.2 Strengths and limitations of the evidence

The reporting of non-CRC diagnoses was sporadic in the reviewed papers and incomplete in many cases, with just specific diagnoses reported (often just GI) therefore cases outside the GI tract may not have been declared. Nonetheless, the diagnosis is based on a highly sensitive test (colonoscopy) so, where stated, the result is likely to be accurate. Adherence to referral target data was poorly reported and is liable to bias depending on the year it was reported from. It was difficult to extract meaningful data from reports hence comparative analysis was not possible.

3.5.2.3 Consistency of findings

This is the first systematic review to report on non-CRC diagnoses or adherence to operational targets in the lower GI TWW cohort. Adherence data is collected nationally and demonstrates poorer compliance with targets related to treatment initiation within 62 days compared with the 14-day initial target referral out outset. The limited data in this review is consistent with these findings.

3.5.3 Survival

Most studies did not include patient follow up and hence did not provide survival data. However, the limited data indicates there was no difference observed between the survival in the TWW and non-TWW cohort, although those presenting as emergencies tended to have a poorer outcome, which correlates with evidence demonstrating they present at a later stage of disease. Data from the National Cancer Intelligence Network demonstrates that there are only marginal differences in one-year survival between patients presenting through any outpatient referral pathway and the TWW. In the TWW cohort around 82% survive one year compared with around 80% one year survival in non-TWW GP referrals (341). This suggests a longer time from referral to diagnosis and subsequent treatment as seen in the non-TWW cohort does not affect survival. However, the impact of morbidity must not

be underestimated: delays to treatment in patients with known malignancy has a significant psychological impact despite no effect on cancer survival.

3.5.3.1 Strengths and limitations of the evidence

The evidence of survival was sparse therefore no firm conclusions can be made on survival in the TWW lower GI cohort from this systematic review.

3.5.4 Impact of the results

The results provoke a reconsideration of the benefits of the TWW pathway for CRC Although there is no defined cancer detection target, there has been a series of DOH proposals: the now archived NHS Cancer plan, published in 2000 declared an aspiration that '90% of cancer be diagnosed via the TWW pathway'(144). This indicates that the pathway is falling short of expectation with a low cancer conversion rate and majority of patients diagnosed with CRC via a non-TWW route, but this is mirrored when looking at nationally collected data on cancer conversion rate via TWW. Conversely the vast majority of TWW patients do not have cancer at all. Latest results from the National Cancer Registration and Analysis Service from 2016 report conclude that around 31% of all CRCs are diagnosed via the TWW route, compared with 10% via screening, 23% via non-TWW GP referral and 23% as emergency presentations (282). This raises questions about the justification for resource prioritisation to TWW patients. A precedent for re-evaluation and change to the current system should be considered.

Key to the difficulties in diagnosing CRC is the poor predictive value of lower GI symptoms for CRC: they are often vague with a large cross over with symptoms of benign disease, therefore a referral system heavily based on symptoms is highly likely to result in investigation of patients with benign disease or no disease at all. So-called 'soft' symptoms such as abdominal pain, constipation and diarrhoea with PPVs of around 1% (146). Anaemia is a more useful referral criterion with a higher PPV. A case –control study found the PPV of a Haemoglobin <11 in men and < 10 in women for CRC was 13.3% and 7.7% respectively in the over 60s age group (344). This is supported by anaemia sub-group in this review where a higher incidence of cancer was found.

In 2015 the NICE referral guidelines were revised with inclusion of symptoms with PPV of 3% or more for cancer, a reduction in PPV from 5%. As a future study it would be useful to compare outcome data before and after this change was implemented to observe its effect on cancer conversion rate. Khong and Peacock both analysed data shortly after a Government awareness campaign. It is interesting to note that this did not significantly affect cancer conversion rate but did increase the total number of TWW referrals(295, 300), suggesting that heightened awareness has a bigger effect on the 'worried well', without targeting those with disease.

Although not a defined outcome measure in this review, it is well documented that adherence to the referral criteria by GPs is poor. The reasons for this are multifactorial. Firstly, the vast majority of what GPs deal with daily is benign disease. A full time UK GP will see a new cancer approximately once a month (any cancer), they therefore become expert in diagnosing what is not cancer, as opposed to what is cancer (146). In addition up to 10% of general practice consultations are due to abdominal pain and most will have a benign cause (103). The National Audit Office reported in 2010 a more than fourfold variation in referral rates across England and Wales. Although some of this can be explained by variations in proportions of elderly in the population, and the impact of social differences such as smoking and alcohol intake, this does not fully explain the referral behaviour (345).

GPs face an enormous challenge in balancing their referral behaviour with the knowledge that they are judged predominantly on their pick-up of cancer, with little credit or recognition for management of benign disease. There is, perhaps, a culture of criticising the GP at secondary care level for inappropriate referral behaviour, but well-recognised are the non-specific symptomatology that many cancers cause that can often be attributed to benign disease. It is not surprising that primary care fast-track referrals are more readily initiated for cancers with more specific presentations e.g. breast or testicular (346). Some GPs have welcomed the idea of feedback on their referral patterns and 'hit rates' in terms of cases referred that have malignancy. Some feel this will have a positive influence on their decision making,

but conversely, others may feel even more pressure to ‘get it right’ leading to high-pressured decision making which neither benefits the patient nor the GP. Certainly more cohesion between primary and secondary care is likely to beneficial.

Lack of an alternative option for patients with symptoms to be seen rapidly, encourages TWW referrals from primary care. Perhaps if the time-frame for urgent and routine appointments were more reliable, allowing the GP reassurance that their patient would be seen within a set time-frame, they might be less inclined to use the TWW pathway when a patient does not entirely meet the criteria. Before the TWW process was introduced, it was clear that patients were waiting vastly different amounts of time for investigations depending on the pressure facing each NHS trust and regression to this system would be wrong. However the current targets are without good quality supportive data to justify them and this review does not support an improved stage at diagnosis.

Undoubtedly there is psychological morbidity associated with waiting for investigations for suspected cancer, however a pragmatic approach must be taken in the financially strained NHS. An all-encompassing referral pathway that all patients with lower GI symptoms can be referred through, e.g. six week-pathway, with subsequent targets for diagnosis and treatment has been suggested by some at both primary and secondary care level. It is unclear whether this would be feasible or, conversely, put more strain on the system by holding every patient to a tight time-frame. A recent study by Redaniel et al gauged clinicians’ opinion on the CRC TWW pathway: many acknowledged they used the pathway knowing some patients did not quite meet the referral criteria, but they had clinical concerns about them and felt reassured they would be seen within two weeks (149). This is likely to be a common scenario that would be addressed by a one-target referral pathway regardless of symptom.

The move by the DOH to encourage increased direct access tests for GPs may be a positive step in the diagnosis of cancer at an earlier stage or conversely corroborate a suspected benign diagnosis, but this will need monitoring to evaluate whether this is just simply increasing a trend towards more investigations in the well or whether it constitutes a less convoluted diagnostic journey for those with malignancy (347).

At secondary care level, once a referral is made, investigation pathways are variable and devised to balance cancer pick-up rates with logistical feasibility given the demands the system places on a hospital trust. Several NHS trusts who collected data included in this review used, or suggested tailoring, the referral pathway to a straight-to-test model or alternatively phone-based consultations. Innovative ways of meeting the heavy demands due to rising referrals numbers are important. Their results demonstrated success with a potential cost advantage by avoiding additional steps in the pathway to diagnosis. There is no suggestion from this review that these pathways improved cancer conversion rates, but they did not seem to miss cancer either. A large-scale case control study may provide more evidence for the feasibility of this model.

The introduction of an objective laboratory test in risk stratifying a patient with symptoms has the potential to increase both sensitivity and specificity of referrals in the future and would mark a shift away from a predominantly symptom-based referral system. Although FOBT was introduced to the 2015 NICE lower GI suspected cancer referral criteria, there is currently no robust evidence for its use in the symptomatic population, particularly using the guaiac faecal test which is still in use. An update to the guidelines in 2017 introduced FIT testing to the referral criteria for patients with 'unexplained symptoms'. Although there is little data on its use in the symptomatic population, it is a rapidly expanding field of research. As a screening tool, there is already a large body of evidence that FIT is superior to the standard guaiac faecal occult blood test (134, 348, 349), although implementation across the UK is still in planning stages. Currently FIT seems to perform poorly for detecting adenomas, but it is likely in the next few years that there will be an increased understanding of the role of FIT in CRC detection outside the screening setting with probable key role in risk stratifying TWW referrals for suspected CRC, with a negative FIT as a rule out test. Widlak et al looked at 430 patients from the TWW cohort using FIT and FCP in combination to assess positive and negative predictive values for CRC. NPV were high at 99%, but this did not alter whether FIT was used in combination with FCP, or alone (350).

Also of increasing interest, and implemented in the United States is multitarget stool DNA testing. Currently one commercial test, called Cologuard is available and works to detect ten DNA-based markers and one human haemoglobin biomarker in stool. Results are combined to produce a diagnostic algorithm which equates to a positive or negative test (101). Results from one blinded screening study were impressive, with sensitivity of 92.3% (351) and results were similar from a second study with sensitivity and specificity data seen at 100% and 93% respectively, although numbers were small (352). Currently this test is not available within the NHS but may hold promise in the future.

3.5.5 Strengths and limitations of the review

This is the first systematic review of all available evidence on key outcomes of the lower GI TWW pathway. There are several limitations to be acknowledged. Firstly, the restrictions created by lack of data reporting in some studies limits the comprehensiveness of evidence and precision of the review findings. However, exclusion of these studies would have led to data gaps. In addition, where reported, demographic data was largely consistent meaning it is reasonable to assume in a similar unrestricted TWW cohort in another area the population should be generally comparable. Due to slight deviations in TWW protocol papers from Scotland and Northern Ireland were excluded at the full paper stage and therefore the results of this review may not be generalisable to the whole UK population. As already acknowledged the heterogeneity of the studies was moderate to high when results were pooled in the meta-analysis.

3.5.6 Future research recommendations

Future research recommendations according to the PICOTS (Population, Intervention, Comparator, Outcome, Timing, and Setting) framework along with limitations in evidence are provided below:

Population and setting

Limitations: There is a paucity of well conducted studies on the TWW cohort with complete demographic data including race, BMI, smoking status and other important variables.

Recommendations: Evidence from well conducted prospective cohort studies with comprehensive demographic data that can be matched against key outcomes such as cancer conversion rate and stage, would be welcomed and provide crucial evidence allowing better risk stratification of the TWW cohort.

Intervention and comparator

Limitations: Although the TWW cohort was clearly defined, most studies lacked a comparator group and those that did provided inconsistent comparator outcome data.

Recommendations: Clearly defined comparators such as urgent or routine referrals are extremely useful when calculating effect of key outcomes in conjunction with the TWW group and facilitate risk of bias assessment better.

Outcome

Limitations: Apart from cancer diagnosis, there was significant variability in outcome reporting e.g. cancer stage was reported as Dukes, TNM, or just as a percentage with metastases making comparisons more difficult. Also polyp detection rate was ambiguous in some cases as it was not clear if polyps or adenomas were counted

Recommendations: The use of accurate and standard methodology for reporting cancer staging and adenoma conversion rate would ensure more valid estimations and comparability.

Timing

Limitation: There was a lack of follow up in the majority of studies meaning survival data was scanty and, where present, was inconsistently reported.

Recommendations: Five-year follow up data would provide robust information on mortality to fill a knowledge deficit in TWW survival.

Policy implications

The findings of this review will help to inform those implementing the TWW pathway of its effect on key performance measures. It also serves to provide feedback to both primary and secondary care of the possible outcomes for their patients who meet TWW criteria.

3.6 Conclusion

This systematic review is the first to summarise comprehensively, key outcome data for the TWW lower GI referral pathway in a large subset of patients. Results indicate that cancer conversion rate is low with no difference in stage at diagnosis using this pathway compared with other routes to diagnosis. Clearly the Government's 'cancer-plan' holds focus on not missing cancer, and picking up more cases. This is not supported by the results of this review. A referral system heavily weighted on symptoms has a low PPV detecting malignancy. Solutions to the problem of balancing improving cancer detection without over-saturating an extremely stretched system are not easy to devise. Target-driven health services are important, but must be regularly evaluated to ensure they serve their purpose and that costs are justified. Several areas for change have been suggested in this review in order to improve the cancer conversion rate: in particular the wider use of FIT testing in conjunction with symptoms is likely to provide an adjunct to better predict cancer in the symptomatic population in the future.

CHAPTER 4:

Methodological development of volatile organic compound analysis in the detection of colorectal cancer

4.1 Introduction

There are many analytical instruments that are able to detect VOCs. The choice is dependent on a variety of factors including experience, expertise, cost and accessibility. In order to transcend the clinical environment for point of care testing, ideally, they need to be portable and easy and cheap to use. They must also be highly sensitive and selective, as VOCs are likely to contain a variable mix of compounds with a variety of molecular weights (156). GC-MS is often seen to be the gold standard analytical technique, as it provides chemical separation and isolation allowing the observer to potentially pinpoint the required chemical (156). However, GC-MS has many practical drawbacks when being used in a clinical environment. For example, it is large and immobile, and requires specialist training for use and interpretation of results (170, 232). FAIMS collects data by detecting ions based on their mobility. This technique has been selected for the majority of VOC experimental analysis in this thesis. FAIMS was chosen due substantial experience gained by previous FAMISHED research groups, which demonstrated the diagnostic ability of FAIMS in a variety of diseases (154, 236). Furthermore, FAIMS demonstrates high levels of sensitivity in ion detection and because of its portability, has potential for use within the clinical setting as a point of care diagnostic tool. However, GC-IMS was used for a subset of patients as an emerging diagnostic tool with exciting potential, in order to make comparisons with the FAIMS results and also to capture chemical information in an attempt to establish the identity of the chemicals within the urine samples.

In this chapter the analytical techniques used in this thesis will be described in detail. In addition, a series of experiments that were carried out for methodological development purposes in order to optimise sampling technique and reduce factors that may confound results are reported.

4.2 Analytical techniques

4.2.1 FAIMS

FAIMS is also known as differential mobility spectrometry (DMS). It is a gas detection technology that identifies chemical ions based on their mobility under high electrical fields. Sample analysis involves a number of stages (figure 4.1) (353).

FAIMS can detect volatiles from aqueous, solid and gaseous sources, which means it can be used to analyse a wide variety of bodily products. Sample preparation will depend on the type of sample you are analysing and the user's preference. Throughout this thesis the Lonestar (Owlstone, UK) instrument was used. There have been several Lonestar software upgrades during the time period of the experimental work. Each upgrade is specified by version number (earliest version is 1 and the latest used was 4.1), the outer body of the instrument has not changed between experiments nor the computer element. The version used for each experiment is denoted within the corresponding method.

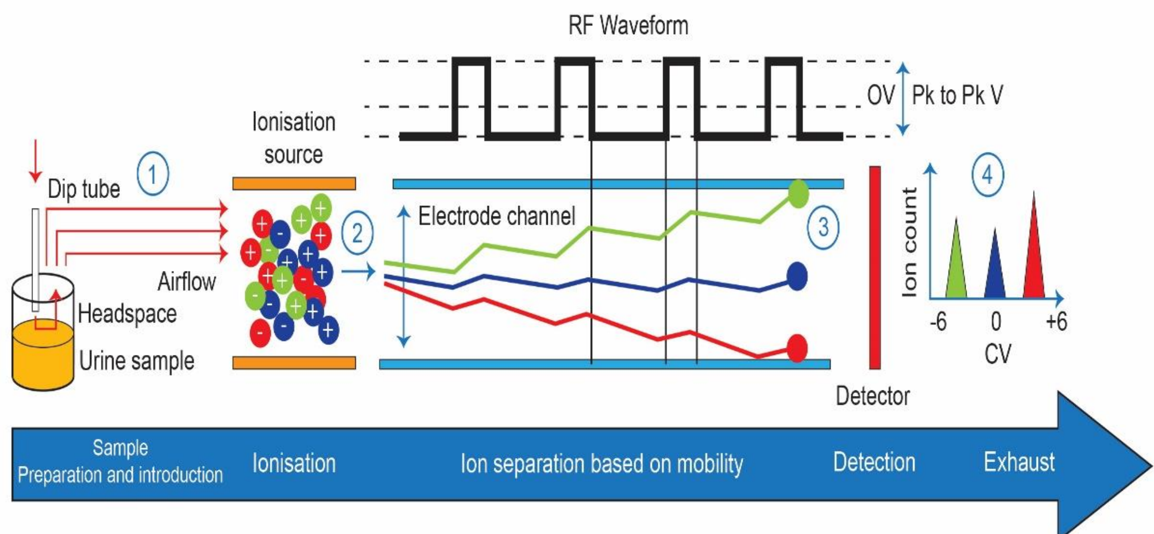
Figure 4.1 Schematic of the FAIMS analysis process *from sample introduction to ion count detection. (Schematic adapted from Owlstone, UK manual with permission(353))*

1. Vapour phase – the sample is introduced via a carrier gas to the ionisation region. Here components are ionised via a charge transfer process or by direct ionisation. Both positive and negative ions are formed.

2. The ion cloud enters the electrode channel, here a radiofrequency (RF) waveform is applied, creating a varying electric field under which the ions follow different trajectories depending on their mobility.

3. A direct current (DC) voltage, called a compensation voltage (CV) is swept across the electrode channel shifting the trajectories to the path of the detector allowing more ions to reach it. The detector then recognises both positive and negative ions.

4. The number of ions detected represents the concentration of chemical in the sample.



Chemical detection using FAIMS

The Lonestar requires a carrier gas to drive ions through the electrode channel to the detector plate, this initiates ionisation. The proportion of ions that make it to the detector increases with increasing flow, improving sensitivity, but it is essential that the carrier gas is free from impurities. Ion detection is based on detecting varying

mobility of ions based on their mass to charge ratio via the following process (170, 353):

Ionisation

An ionisation source is required. In this case radioactive nickel-63 was used. Ni-63 undergoes beta decay which generates energetic electrons that interact with the carrier gas to form stable +ve and –ve intermediate ions. These give rise to reactive ion peaks (RIP) in the positive and negative FAIMS spectra (figures 4.4 and 4.5). The RIP ions transfer their charge to neutral molecules through collisions. The likelihood of ionisation is determined by the analytes affinity towards protons and electrons.

Mobility

Ions in air under an electric field will move at a constant velocity proportional to the electric field, this denotes their mobility. In FAIMS, when the ions enter the electrode channel, applied radio frequency (RF) voltages create oscillating areas of high and low velocity electric fields as the ions are moving through the channel. Depending on the ion's mobility, the trajectory of the ion moving through the channel will change as the high and low electric fields are applied. The composition of a chemical ion will directly affect its mobility through the oscillating electric field, and other factors such as humidity, temperature and pressure all have an effect.

Compensation Voltage

Ions of different mobilities travel along the electrode channel, some will not arrive at the detector as they are destroyed by colliding with an electrode. To increase the proportion of ions reaching the detector, a compensation voltage (CV) is applied between the top and bottom electrode. This serves to re-centre some ions allowing them to hit the detector, it also produces a CV spectrum. Ion mobility can be demonstrated as a compensation voltage at a set electric field.

Factors affecting sampling

Several sampling factors will affect the concentration of vapour analysed:

a) The pre-concentration time is key as build-up of vapour concentration in the headspace increases with time until the headspace is saturated.

b) Sample temperature has a direct effect on altering evaporation rates and the time taken for the headspace to reach saturation. The more volatile chemicals will evaporate more readily. The volatility of analytes can be reviewed by looking at their vapour pressure, where the pressure at which the gas phase of the analyte is in equilibrium with the liquid phase. Increasing the temperature may help to increase the rate of evaporation of a sample. If the chemical tested for is known, then the sampling temperature can be set to optimise detection, however previous experimental work performed found temperatures over 40^0C forced water into the system which disrupted results.

c) It is also important to recognise the effect of analyte solubility on chemical detection. In the case of biological sampling the chemical profile is yet unknown, the more soluble the analyte, the less likely it is to leave the urine.

d) pH also effects analyte solubility, increased acidity can result in more of the analyte being in vapour form, but with other solutions the reverse is true. Again, if the chemical in question is known, the pH of the solution can be altered to maximise headspace vapour concentration (354). In all experiments in this thesis the pH was not altered.

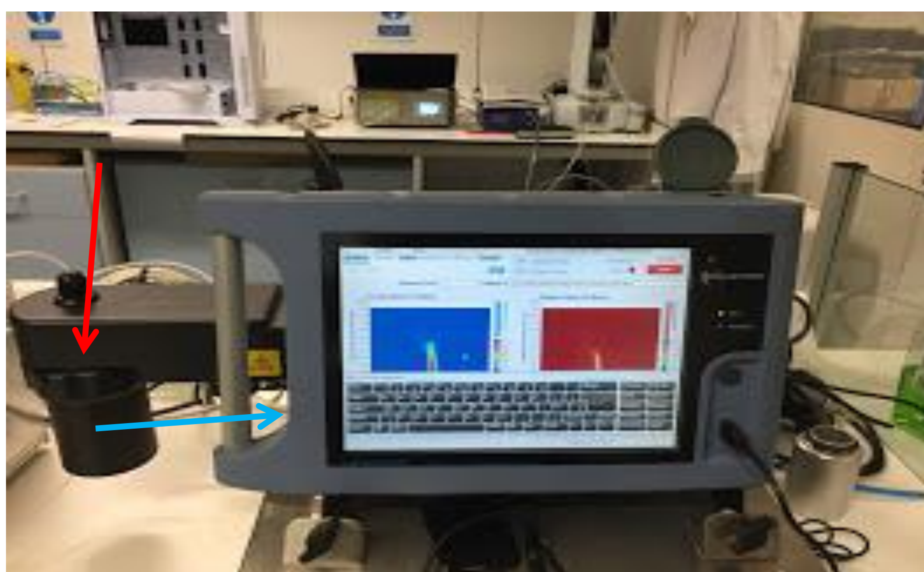
Manual headspace sampling

The At-Line Sampling Module (ATLAS, Owlstone, UK) was designed to be combined with the Lonestar system to ensure that headspace sampling is simple and reproducible. The headspace refers to the compounds in the gas phase above the liquid sample within the vial and this is where the VOCs are extracted from. Sampling occurs as follows (354).

1. A 10ml headspace glass vial (Thames Restek, UK) containing 5mls of thawed sample is placed in the ATLAS sampler. It is important the dip tube (which collects the headspace) does not come into contact with the liquid sample.

2. It is held at a constant temperature of 40 ^{0}C whilst in the sampling device
3. Chemicals from the urine sample desorb and move into the headspace of the vial.
4. A dip tube is lowered into the headspace.
5. A constant flow of clean, dry air (purge flow) is passed through the headspace and into the Lonestar, sweeping the analytes into the detector using a dynamic sampling method.
6. Standard Lonestar settings were: Flow rate 2000 mL/m, 0.978 pressure barg.

***Figure 4.2 Picture of the ATLAS sampling module** (red arrow) which introduces sample headspace to the Lonestar device (gas flow shown by blue arrow)*



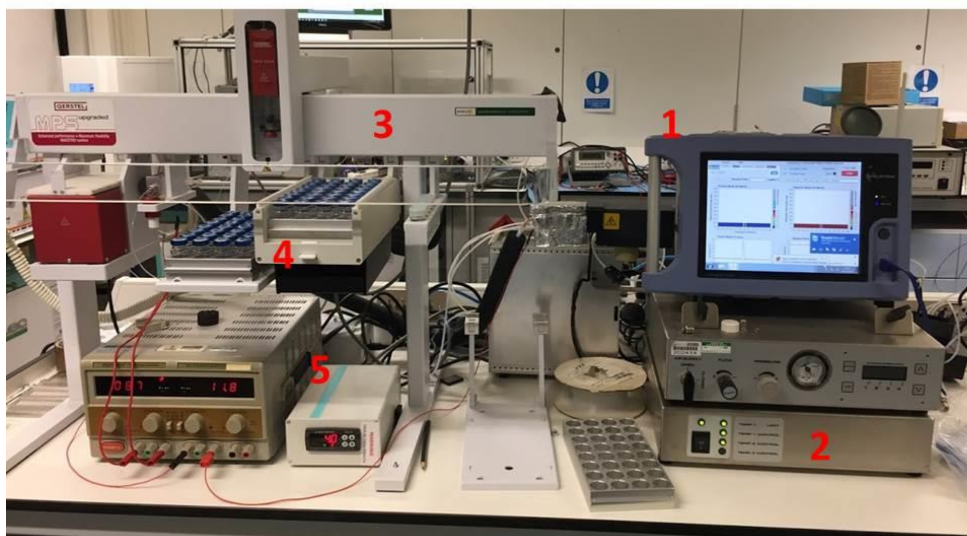
Auto sampling

There is no commercially available auto sampling system produced by Owlstone for use with the Lonestar instrument, therefore the Biosensors team at Warwick University developed a method to connect a modified Gerstal MPS to the Lonestar to improve speed and consistency of sampling.

Sampling method:

1. Once defrosted, 3ml samples were aliquoted out into 10mL glass vials suitable for use with an auto sampler. After the samples were aliquoted, each vial was sealed using the crimp cap and septa (Chromacol).
2. Sample 'blanks' of air were sealed using the same method.
3. Both the air blanks and the samples were then placed into the auto sampler trays. The auto sampler tray for the urine samples was set to 40^{0}C. Once the urine samples were placed in this tray their temperature was allowed to equilibrate for 30 minutes prior to sampling.
4. Flow rates for the Lonestar were controlled by the MFC, settings were determined by earlier experimental work. The settings were as follows: 300mL/min of air to the sample and 1700 mL/min make up flow. The set-up of the auto sampler was bespoke, allowing the sample air flow to move through the auto sampler with a purge tool directly into the sample vial, then back out and into the Lonestar by use of heated transfer lines.
5. Prior to the sample data being obtained, both the samples and the air blanks were heated to 40 ^{0}C and agitated. Air blanks were heated and agitated for 90 seconds, whilst the chilled samples were heated and agitated for 8 minutes. This was automated by auto-sampler software.

Figure 4.3 Picture of the Lonestar FAIMS instrument together with auto sampler*. 1. Lonestar FAIMS display screen 2. ATLAS Pneumatic control box and heater control box 3. Auto sampler 4. Cooling tray containing sample vials 5. Chiller control.*



4.2.2 FAIMS matrix

The Lonestar FAIMS is set up to scan between 0 and 100% dispersion field in 51 steps and a compensation voltage of between -6 V and +6 V in 512 steps. The chemical information gathered is termed a matrix and represents 52,224 data points for both positive and negative ions collectively. Multiple matrices can be gathered for each sample, by performing multiple sample runs where the process is repeated. A matrix is most frequently presented as a 2D colourplot: here colour contours represent data on chemical composition forming a unique spectra. This is dictated by the ion current at each electric field/compensation voltage point and are constructed using a bespoke software program (figure 4.4). A colourplot should have two peaks: the reactive ion peak (RIP) and product ion peak (PIP). The RIP demonstrates the background ion detection present in any environment, as even after cleaning there will still be some contamination in the system, but as this is constant it does not interfere with analysis. The PIP represents information collected from sample analysis. To further demonstrate what the colourplot

represents the same information is shown as a 3D waterfall plot in figure 4.5. Files can be saved within the Lonestar computer, or exported for analysis.

Figure 4.4 Lonestar colourplot showing a dispersion field matrix. *Orange circle represents the Reactive ion peak (RIP) due to 'background noise', green circle represents product ion peak (PIP)representing information collected from the specific sample*

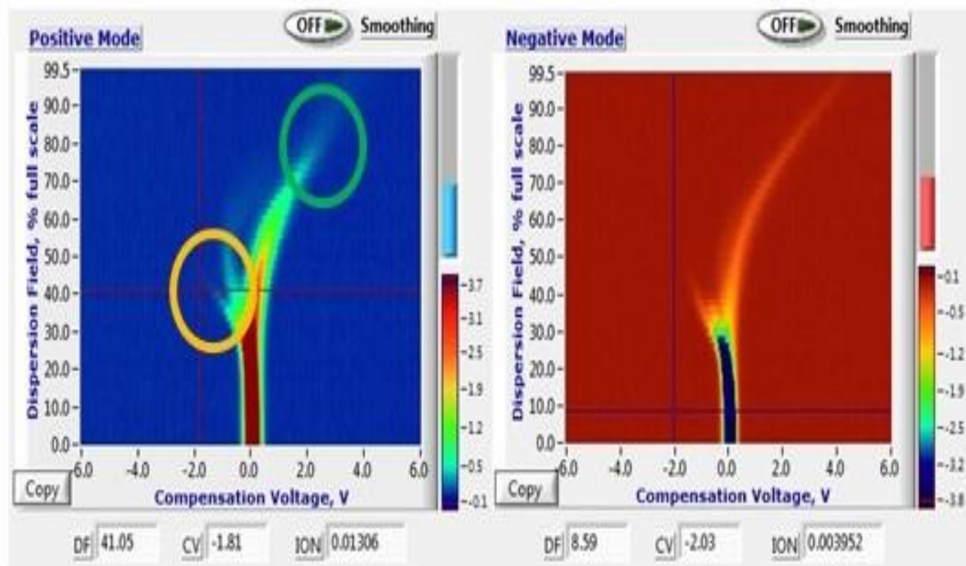
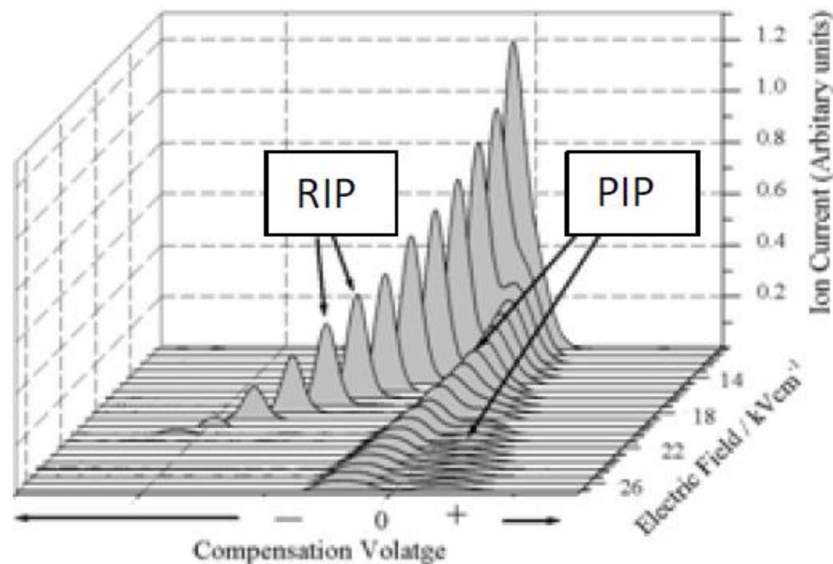


Figure 4.5 Lonestar dispersion field matrix demonstrated by a waterfall plot *showing the same data points as figure 4.4. The z axis is the ion current. (image used with permission from Owlstone, UK)*



In order for this method of chemical detection to be successful in testing biological samples for disease it requires several factors:

1. The disease in question produces a specific set of chemical changes in urine that are seen consistently when the disease is present.
2. This unique chemical fingerprint must be stable within the urine sample under collection and storage conditions.
3. The chemicals must be consistently present in high enough concentrations to allow detection by the Lonestar.

4.2.3 Statistical analysis

Analysis is performed using machine learning methods. This method utilises a computer-devised algorithm that can make predictions on unknown datasets comprised of thousands of data points that are created following the analysis of a single urine sample. Firstly, the algorithm builds a model based on a training set of data where the output, or diagnosis, is known. Once this model is constructed, datasets of unknown outcome can be fed into the algorithm and a diagnostic prediction can be made.

Machine learning methods of data analysis for disease detection is a new, but rapidly expanding area. Multiple data analysis pipelines have been developed, and continue to be advanced in order to improve data classification and disease detection. The three methods used in this analytical work are now described:

Method one

Statistical analysis of data obtained through Lonestar was processed in a custom LabVIEW program (Ver 2012, National Instruments, USA) using Fisher Discriminant Analysis (FDA). FDA assesses the separation in the data between classes of samples. The ion count data for each sample was wavelet transformed to identify specific signals and reduce noise. FDA then identifies a projection of this data onto a one dimensional subspace which maximises the separation between classes. Varying a threshold across this subspace allows Receiver Operating Characteristic (ROC) curve analysis of the resulting FDA classifier to assess the ability of the Lonestar data to discriminate between classes.

Method two

The raw Lonestar data are analysed through a four-step process. This is used as part of a cross-validation analysis, which makes efficient use of the number of samples in the study.

1. **Wavelet transformation**. Raw Lonestar data are known to have a lot of correlation between the different variables, so a wavelet transformation is used to exploit this. This data compression step reduces the dimensionality of the data, helping extract the highly correlated patterns within the data set.
2. **Feature selection.** Mann-Whitney-Wilcoxon rank-sum test is applied to each of the features in turn. This assigns a *p-value* to each feature which can be used to rank variables by how much information they provide on the desired outcome. The top 'n' features are selected and passed to the classifier. Within the cross-validation, only the training data are used in each fold to select these features.
3. **Classifier training**. The data for the selected features are passed into the classifier algorithm (Stochastic Gradient Boosting) the classifier is then

used to predict the data points not used in training the classifier (i.e. the current fold)

4. **ROC Curve**. The classifier gives each sample a probability indicating how likely it is to be a case rather than a control. All the scores from the cross validation are put together to produce a ROC curve.

Method 3

The latest data analysis technique follows on from method one, with data analysed using R studio (R Foundation for Statistical Computing, Vienna, Austria). The analysis pipeline involves the following steps, once again being used inside a cross-validation.:

1. **Data input** – The raw data is collected from the 2D matrix
2. **Aligning and cropping** – Large amounts of irrelevant data is extracted prior to analysis
3. **2D Discrete Wavelet Transformation** – Allows extraction of subtle chemical signals hidden within a much larger signal. This is conducted using the Wavethresh package in R studio.
4. **Feature Selection/ Exclusion** – Applies a simple threshold on the standard deviation of the feature e.g. remove wavelet features where the variance is ~0 (as these are expected to be uninformative). Then using Wilcoxon rank-sum test each remaining feature in turn, keeping only those with the most significant p-values.
5. **Principle component analysis (PCA) –** a further data compression step, to de-correlate the remaining features.
6. **Classification model training** – Five classification models are used:

   a) Sparse logistic regression.

   b) Random Forest.

   c) Gaussian process classifier.

   d) Support vector machine.

   e) Neural network.

   Each model is compared to find the most accurate for the specific data analysed, this may change depending on each dataset. Data are fed into the

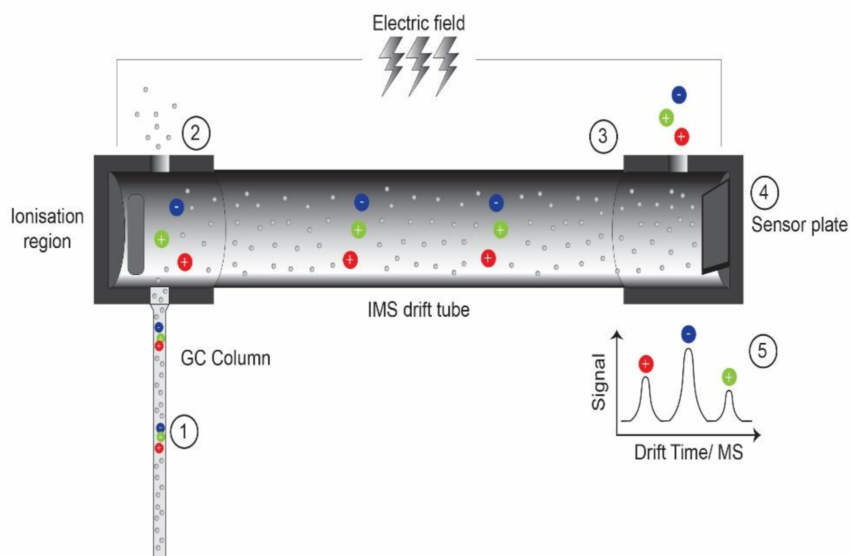
pipeline using samples where the outcome is known. This serves to 'train the algorithm' in order that it may recognise the same data patterns if they present again in another sample e.g. CRC. Once trained, the algorithm can be applied to unknown samples.

7. **Balance the data** – If the two comparator groups have uneven numbers the data needs to be balanced. This is achieved using Synthetic Minority Over Sampling Technique (SMOTE) to achieve unbiased balancing.
8. **ROC analysis –** classification data is presented as a ROC curve
9. **Performance metrics** – Information gathered from this analysis pipeline are Area Under the Curve (AUC), Sensitivity, Specificity, PPV and NPV.

4.3 GC-IMS analysis

Gas-capillary column ion mobility spectrometer (GC-IMS) is a gas column coupled to an Ion Mobility Spectrometer and this has been used in a sub set of patients in this thesis (Chapter Five). It involves a two-stage analytical process; firstly separation occurs by GC and secondly detection via IMS (355). Samples are separated before they reach the sensor, thus theoretically improving chemical detection and, in theory, allowing identification of specific composite chemicals. A schematic of the instrument is shown in figure 4.6. The instrument used in this thesis is the GC-IMS Silox (IMSPEX diagnostics Ltd, South Wales, UK). It has been used in industry for several years, in particular in the food industry for freshness quality control and has theoretical advantages when applied to the medical field for detection of disease. Firstly it has a very fast run time and secondly it is low maintenance with low running costs and is portable, unlike the GC-MS (176, 178, 356, 357).

Figure 4.6 Schematic of the separation process and ion detection using GC-IMS. *1. Sample passes through the gas column where initial separation occurs. 2. The discrete compounds are consecutively fed into the ionisation chamber where ionisation occurs 3. Ions pass through the drift tube at varying speeds dependent on their mobility 4. Ions hit the sensor plate and are detected 5. Ion peaks are calculated based on drift time.*



Data output is twofold: GC gives peaks representing retention time as they pass through the column. This is coupled with IMS data, based on the mobility of the ions as they pass through the drift tube and hit the sensor as represented in figure 4.7. The culmination of this two-phase analysis is illustrated in an IMS chromatogram which is represented by millions of data points in a heatmap (figure 4.8). These data points are subject to very similar statistical analysis as is applied to the Lonestar data in method three above.

***Figure 4.7 The two-stage separation of a biological sample using GC-IMS**. Initial separation using a gas column and subsequent separation according to ion mobility (image adapted with permission from Imspex, UK)*

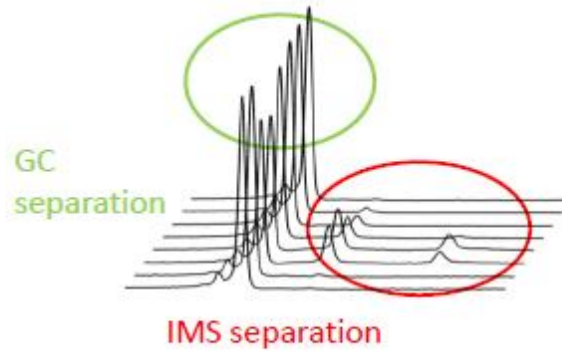
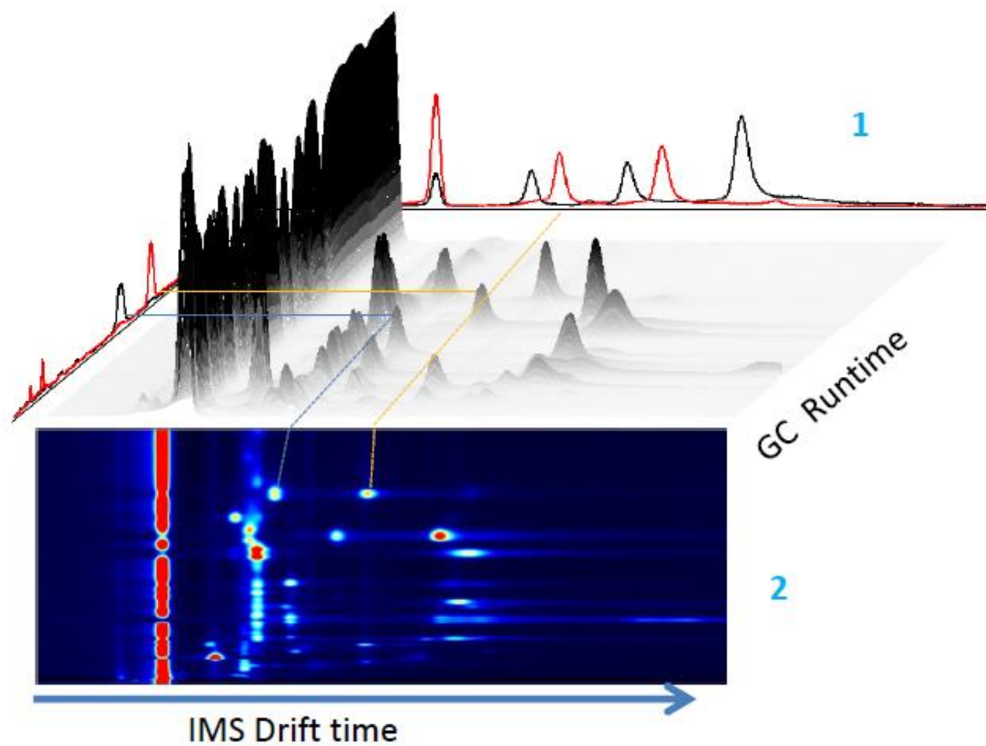


***Figure 4.8. 3-dimensional representation of GC data output with corresponding IMS chromatogram**. 1. Single IMS spectra data is combined with GC run time peaks. 2 Heatmap corresponding to GC-IMS peaks (yellow and blue lines) (Image adapted with permission from Imspex, UK)*



4.3.1 Manual sampling method

1. The urine sample is stored and then thawed as per the method for the Lonestar.
2. A five ml urine sample is aliquoted into a glass vial and sealed with a crimp lid.
3. A 21g needle is attached to the GC-IMS input port.
4. The needle with attached port is inserted into the sample headspace one cm above the urine.
5. The needle is held in place for twenty seconds to allow for vapour aspiration.
6. The total run time is five minutes per sample.
7. The carrier gas flow rate is 150ml/min and sample flow rate through the instrument is 20ml/min.
8. The sample heating was carried out in accordance with manufacturer instructions to a maximal level of 80 ^{0} C.

GC-IMS has only been studied in the detection of disease in a handful of cases (174, 175) including a recent pilot study of patients with respiratory tract infections (176). Results were encouraging but work using this modality is only in the pilot stage.

4.4 Experimental work

The application of FAIMS in the diagnosis of disease is novel and the analysis of biological samples is still in the research phase. In order to understand the optimal sample storage and analysis conditions a series of experiments were conducted on both healthy volunteers and a selection of disease groups. In addition, a number of setbacks were encountered due to technical and practical issues whilst running case and control samples. This led to re-evaluation of the methodology and run settings of the instrument, with consequent method improvement.

These key experiments are described below in chronological order:

4.4.1 Experiment one: urine validation

Data analysis: April 2015

Background: During repeated use for analysis and due to background VOCs present within the lab, it is common for the Lonestar to still show evidence of signal intensity reflected in the plume even after repeated air or 'blank' runs (seen as the RIP). This low level signal intensity, is referred to as 'background noise'. Background noise is expected in any environment that the Lonestar is being used in as there is a constant supply of VOCs present. However, it is crucial that this noise is consistent throughout testing because during analysis a signal that persists and is common to each sample analysed can be removed from data analysis, but if there is fluctuation in the background noise or high intensity signal pick-up then this may well confound results.

A 'run' refers to the flow of headspace from a sample through the Lonestar analysis process. Multiple runs per sample are required in order to ensure that signal pickup is maximised. It is important when deciding on the number of runs to perform per sample, that ion current intensity is not increasing with each sample. The ion current represents the amount of ionic activity passing through the sensor and although this cannot specify which ions are being detected, it potentially correlates with the key targets within the urine sample. When selecting the number of runs, ionic activity needs to be lower on the last run that it was on previous runs, i.e. show a decreasing pattern. This suggests activity has peaked and the strongest ion concentration has been captured.

Method: During February 2015, 2x20ml urine samples were collected from 25 randomly selected medical students at UHCW for use as 'validation samples', to optimise sample run parameters. This group were chosen for validation reasons with the assumption that due to the lack of complex medical history or current pathology their urine would not produce significant signal change in the Lonestar output.

Participants were consented and completed a standardised FAMISHED CRF with basic demographic and medical detail (table 4.1).

Table 4.1. Study characteristics of the urine validation cohort

Female (%)	**Median age (mean)**	**Non-smokers (%)**	**Mean alcohol intake (u/wk)**
15 (60%)	24 (24.8)	25 (100)	6

One was vegetarian, the rest had a normal diet. There were no significant co-morbidities.

Ionic activity per run is shown below in Table 4.2 for each healthy control. Each matrix represents a sample run. The fourth matrix demonstrates a decreasing ionic concentration in the majority of samples meaning that the peak ionic concentration has been captured containing the most intense VOC data.

Table 4.2 Ion concentration data per FAIMS matrix *for each urine validation sample (peak concentrations in red)*

Matrices	**1**	**2**	**3**	**4**	**Matrices**	**1**	**2**	**3**	**4**
Sample No					Sample No				
1	1.261	3.971	4.155	4.116	**14**	0.892	4.009	4.183	4.132
2	1.122	3.654	3.671	3.616	15	1.259	4.14	4.306	4.228
3	1.215	3.629	3.923	3.858	**16**	1.364	4.076	4.287	4.268
4	1.236	3.865	4.01	3.928	**17**	1.367	4.285	4.636	4.711
5	1.112	3.646	3.932	4.173	**18**	1.294	4.559	4.619	4.603
6	1.3	3.834	3.99	3.978	**19**	1.323	4.096	4.273	4.093
7	1.337	4.073	4.187	4.101	**20**	0.999	3.902	4.064	4.073
8	1.027	3.22	3.355	3.297	**21**	1.003	3.907	3.93	3.87
9	1.352	3.865	3.95	3.936	**22**	1.2	4.184	4.244	4.172
10	1.348	3.855	3.982	3.988	**23**	1.13	4.393	4.427	4.337
11	1.15	3.956	4.066	4	**24**	1.15	3.903	4.047	4.018
12	1.322	3.977	4.113	4.087	**25**				
13	1.041	3.861	4.17	4.334					

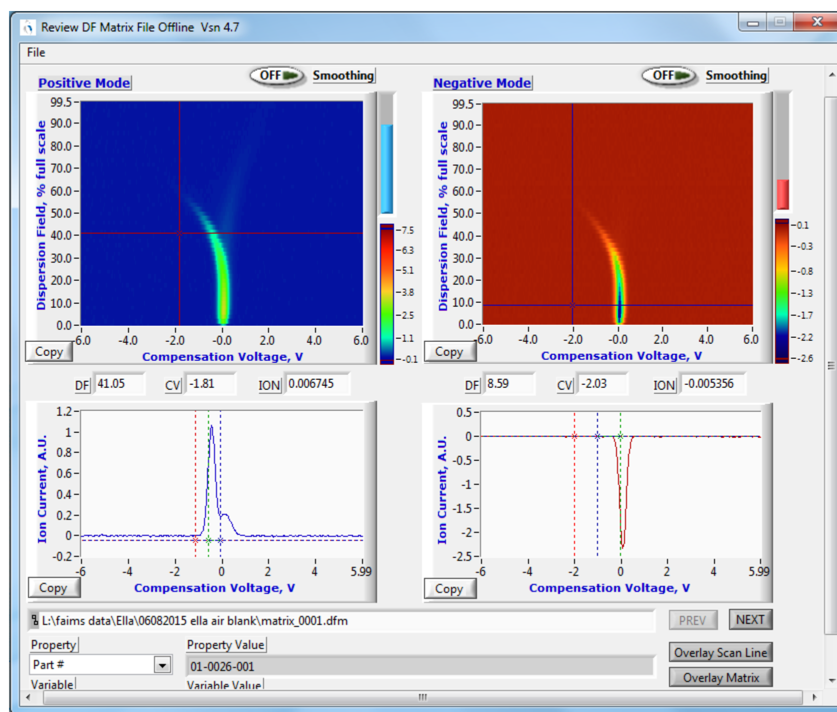
Following a sample run, the Lonestar needs to be 'washed-out' before a second sample can be analysed. This is to ensure no residual information from the previous sample is still present within the Lonsestar instrument. As with sample run number, the number of 'air' or 'blank' runs that need to be performed will vary depending on the sample under analysis. Through previous experimental work ten air runs was found to be optimal, allowing sample washout with a marked decrease in ion concentration. This was confirmed during the urine validation analysis: the Lonestar was adequately 'clean' after ten runs. Each run took 2.5 minutes to complete. Therefore each sample/air cycle was 35 minutes. Sample analysis was performed manually, therefore it was predicted that approximately 12 samples could be run per day (allowing for sample thawing).

4.4.2 Experiment two: CRC and controls

Data analysis: July 2015

88 CRC and 40 IBS sample urines were analysed by manual processing on the Lonestar v1 instrument. Although collection of sample headspace was achieved as expected, following sample processing, it was found that a major 'drift' had occurred on the plume. A drift refers to the shifting of the plume, demonstrating an atypical dispersion field. This signifies an error in sample processing. Therefore any difference in signal change between CRC and IBS sample will not be detected accurately. This may be due to contamination of the system or a fault with the ionisation process. In this case there was a build-up of charge on the Lonestar chip.

Figure 4.9 An abnormal Lonestar plume *in both positive (blue) and negative (red) mode. Here there is poor signal intensity with little variability in ionic capture. See figure 4.4 for comparison.*



Following this technical issue, a replacement chip was inserted resulting in resolution of the error. This unforeseen problem was only detectable once the samples had been analysed. The instrument also had a hardware/software upgrade to Lonestar v3, resulting in improved processing and faster run time, allowing a run to complete in 45 seconds. This, so called, fast scan allows the same amount of data capture as the prior slower version with 15 runs being completed per sample with 30 air runs to ensure adequate cleaning.

4.4.3 Experiment three: sample storage

Data Analysis: August 2015

<u>Background:</u> There is little in the literature about the storage of urine specimens for VOC analysis and whether the storage and thawing process has a significant effect on the number and type of VOCs seen during analysis. This work was led by Esfahani and Sagar with myself as a co-author and is published in PLoS one (358).

The aim was to understand the rate of urinary degradation during long-term storage at -80 degrees.

Method: Patients with type 2 diabetes mellitus were recruited from UHCW between December 2009 and May 2014 as part of the FAMISHED study. Once collected, samples were initially stored at -80^{0}C. 87 samples were analysed using the Lonestar v2 manually. All had T2DM as defined by WHO criteria, exclusions were IBD, IBS, coeliac disease and malignancy.

Results: The analysis showed that total chemical concentration decreases with sample age (figure 4.10), with a particular decline after nine months of storage. The chemical diversity of the samples is defined as the amount of different chemicals in the sample. This was higher in samples of increasing age, but most prominent in samples older than nine months. (figure 4.11).

Figure 4.10 – Scatter plot showing change in total number of urinary VOCs over time *(from December 2009 to May 2014). The graph shows the total amount of chemicals released (number of dots) decreases with age and the concentration also decreases (drop in ion count with time).*

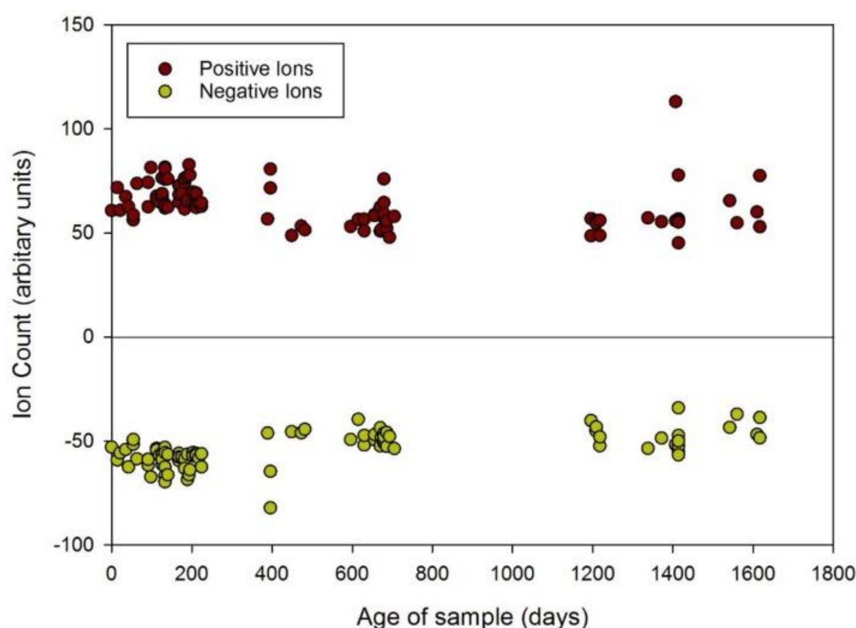
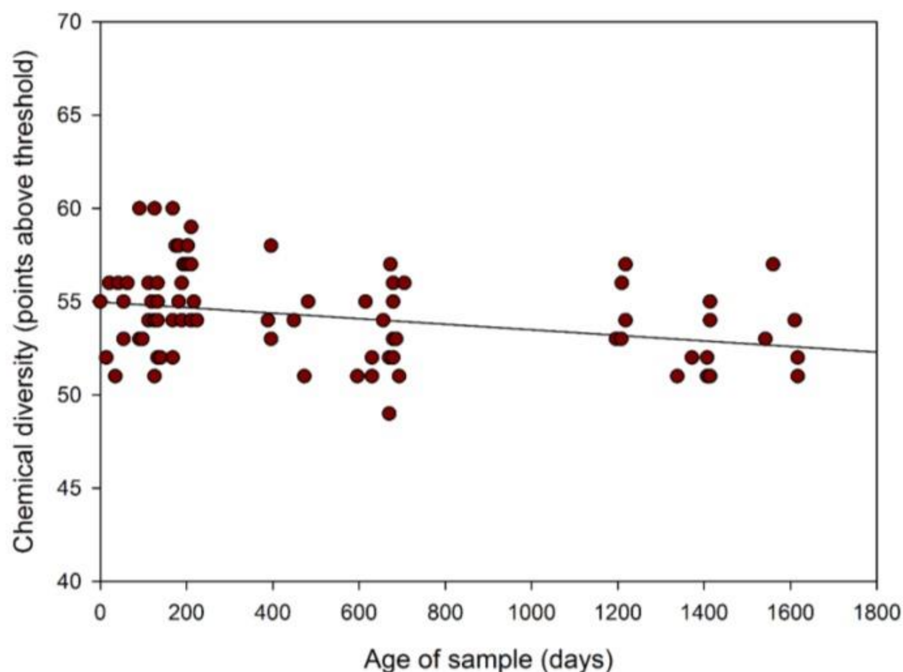


Figure 4.11 Scatter plot showing chemical diversity of urinary VOCs over time *(from December 2009 to May 2014), the red dots represent different chemicals within samples of differing ages, a larger number of dots equates to more chemical diversity. A threshold for chemical detection was set to avoid picking up 'background noise' which represents chemicals within the sample due to the environment. Linear fit emphasises output change.*



In conclusion, with increasing age, even stored at -80^{0}C samples are losing ion count and diversity is changing. This must be taken into account during sample analysis. The changes seem more predominant after nine months of storage. Further experimental work, looking into storage and determining whether the sensitivity of disease detection is affected over time would be extremely helpful. This study tentatively recommends storage of samples for a maximum of 12 months but ideally they should be analysed within nine months. However, the practicalities of sample analysis within nine months need to be considered within the research setting as it often takes a longer period to recruit a satisfactory number of patients.

4.4.4 Experiment four: urine degradation

Data analysis: February 2017

<u>Background:</u> There is no published data on the effect of variation in time at room temperature, prior to urine sample freezing for storage. This is an important

consideration in the clinical setting. Depending on the location of sample collection, the time of day and access to freezer storage, it would be common for samples to be kept at room temperature for varying amounts of times after recruitment.

Through joint experimental work between myself and Dr Michael McFarlane, the effect of room temperature on total ion count and ion variability using the Lonestar instrument was explored. This work is included in Dr McFarlane's MD thesis submitted to the University of Warwick.

Method: 26 UHCW outpatient and endoscopy staff were recruited at random during October 2016. Basic patient information was collected via the standard FAMISHED CRF. 2x 20ml universal container of urine was collected from each patient. Once collected, urine was immediately aliquoted into six universal containers, labelled with the study code and the additional number 000, 012, 024,036,048 or 072 to denote the number of hours the sample would remain at room temperature. Sample 000 was placed immediately into the freezer within the research department at -20^{0} C and transferred to -80^{0}C in tissue bank at UHCW for long term storage within 48 hours of collection. Sample 012 was left at room temperature for 12 hours before freezing. Samples 024,036,048 and 072 were left for 24, 36,48 and 72 hours respectively at room temperature prior to freezing. Lonestar v3.1 with attached auto sampler was used for analysis. All samples were analysed within the same week. Analysis time was 10 minutes which equates to approximately 15 runs per sample. Air runs were completed between samples to ensure no residual sample was present in the instrument. Analysis was completed using Fisher Discriminant Analysis after feature extraction from the raw data (statistical method one). Spearman's Rank correlation coefficient was used to assess the relationship between the two variables.

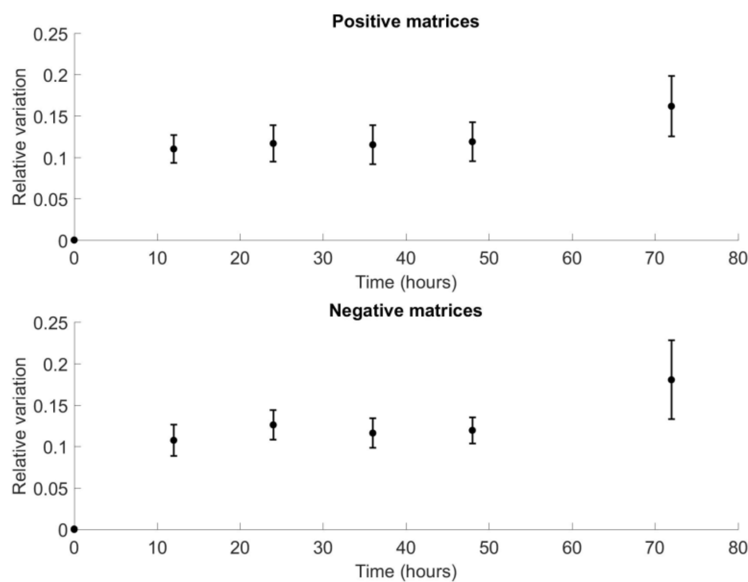
Results: 20 complete samples were analysed. 15 (75%) were female with an average age of 49 years.

Three important observations were seen from this analysis:

1. Variation from baseline. When comparing Lonestar matrices as an average over all samples, signal variation increased with time. There was a slight

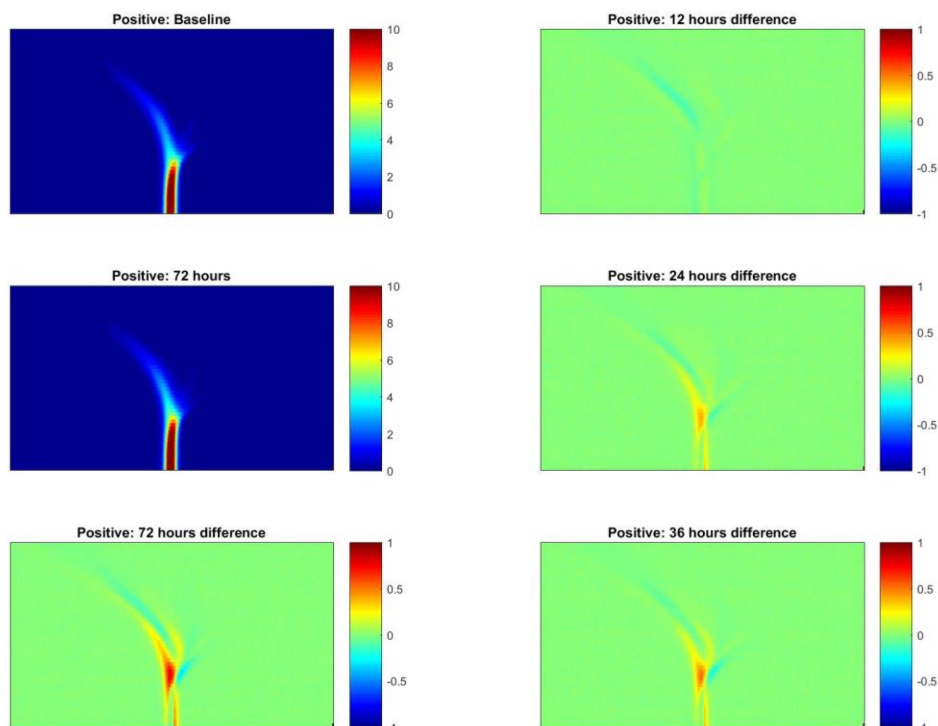
plateau at 12-48 hours with a further rise at 72 hours. This pattern was seen with both positive and negative matrices (figure 4.12). Spearman's rank correlation coefficient was used to test for a monotonic relationship between time and variation from baseline (where an increase in the independent variable is associated with an increase in the dependent variable). Spearman $\dot{\rho}$ = 0.94, p=0.017 for positive matrices and $\dot{\rho}$=0.83, p=0.058 for negative matrices.

Figure 4.12 – Chart of variation of FAIMS matrices, as an average of all samples with increasing time*. Bars show standard error. Relative variation increases with time at room temperature, although it plateaus at 12-48 hours with a particular rise in signal production at 72 hours.*



The greater variation in matrices with increasing time can also be visualised in the plume structure with increasing time (figure 4.13). The same spectra changes with time are observed in the negative matrices too.

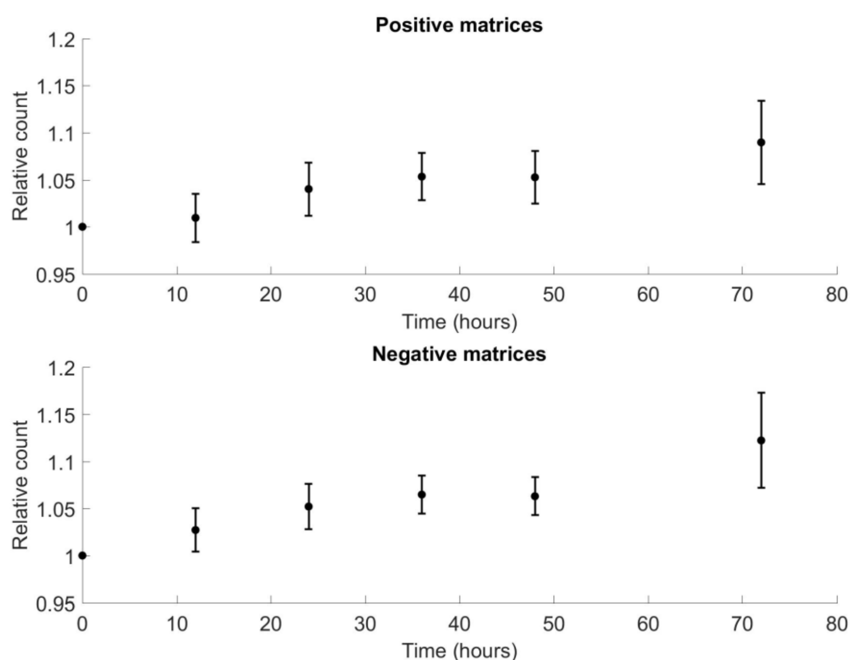
Figure 4.13 Variation in FAIMS matrices as seen by plume structure*. Top left: FAIMS plume at time zero. Middle left: FAIMS spectrum at 72 hours. Bottom left: Difference between spectrum at 72 hours and baseline. Right hand panels show the difference at 12,24 and 36 hours, showing a consistent pattern developing with time*



.

2. Changes in total ion count: There is an increase in total ion count, observed with increasing time at room temperature before freezing (figure 4.14). The average total ion number detected for each patient at each point in time was compared with the average number of ions detected at time zero for that patient. For positive matrices the mean ion count with time gives Spearman ṗ=0.94 p=0.017 and ṗ=0.94 p=0.017 for negative matrices.

Figure 4.14 - Chart of trend in total ion count from mean of FAIMS matrices. *Bars indicate standard error. The top panel shows the positive matrices and the bottom shows the negative ones. There appears to be a positive correlation between relative ion count and time.*



Variation and total ion count with matrix number: When looking at the pattern of degradation during sample analysis with each consecutive matrix or run, the positive matrices showed a consistent monotonic decrease in ion count and increase in variation as each matrix was formed (see figures 4.15a and 4.15b and 4.16a and 4.16b respectively). Conversely, when looking at the negative matrices a transient increase in ion count is observed, with subsequent decaying and an increase in variability. This pattern did not seem to be affected by how long the sample was stored at room temperature, so each pattern was replicated in all samples irrespective of time. This loss of ion count can also be visualised when observing the plume, or Lonestar spectrum structure over subsequent matrices (figure 4.17).

Figure 4.15a – Positive FAIMS matrices: variation in ion count from baseline as a function of matrix number *for each time point, showing decreasing ion count per matrix. Bars show standard deviation.*

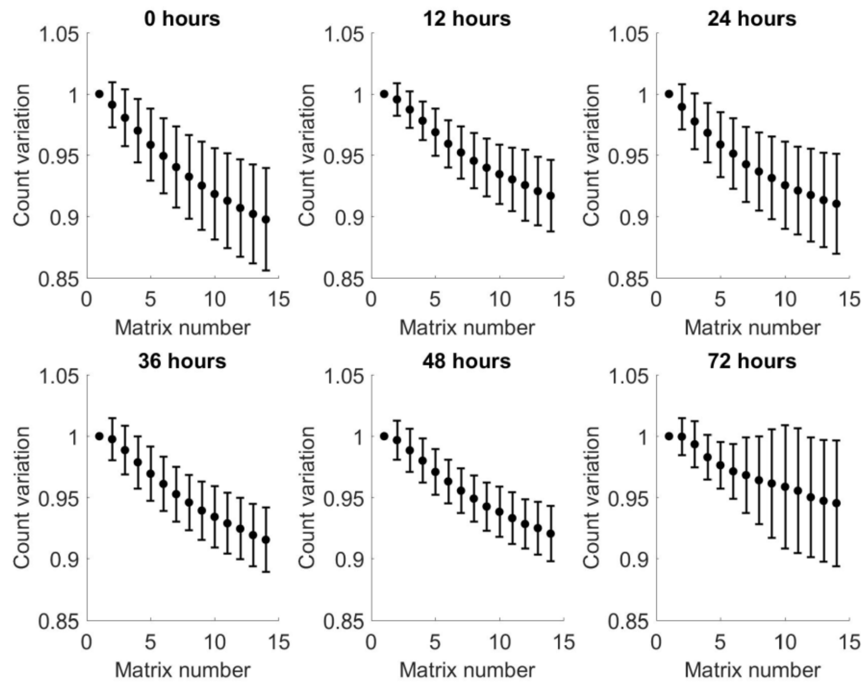


Figure 4.15b - Negative FAIMS matrices: variation in ion count from baseline as a function of matrix number *for each time point, showing decreasing ion count per matrix. Bars show standard deviation.*

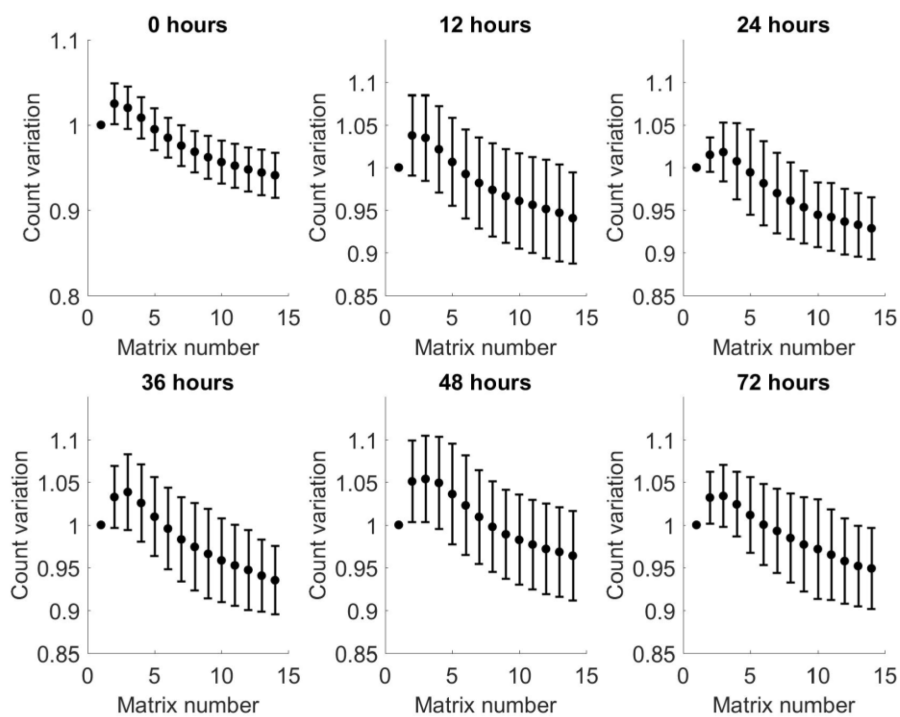
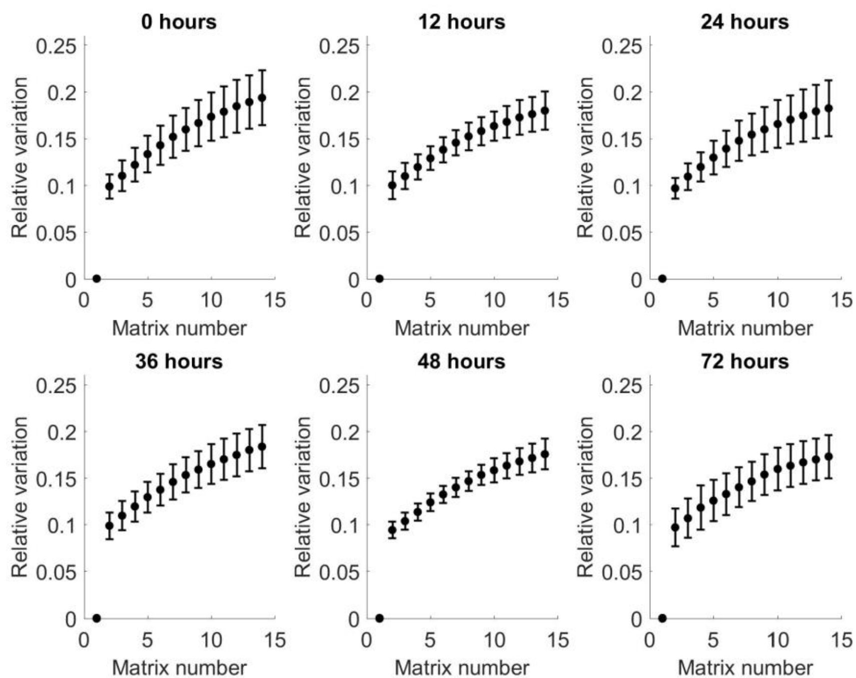


Figure 4.16a - Positive FAIMS matrices: variation from baseline as a function of matrix number *for each time point, showing increasing variation with each matrix. Bars show standard deviation.*



.

Figure 4.16b - Negative FAIMS matrices: variation from baseline as a function of matrix number *for each time point, showing increasing variation with each matrix. Bars show standard deviation*

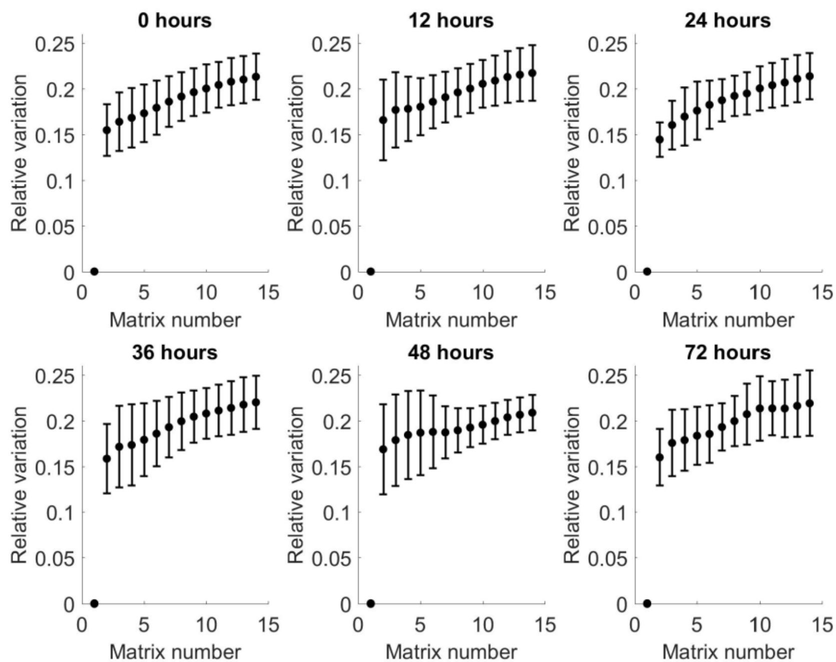
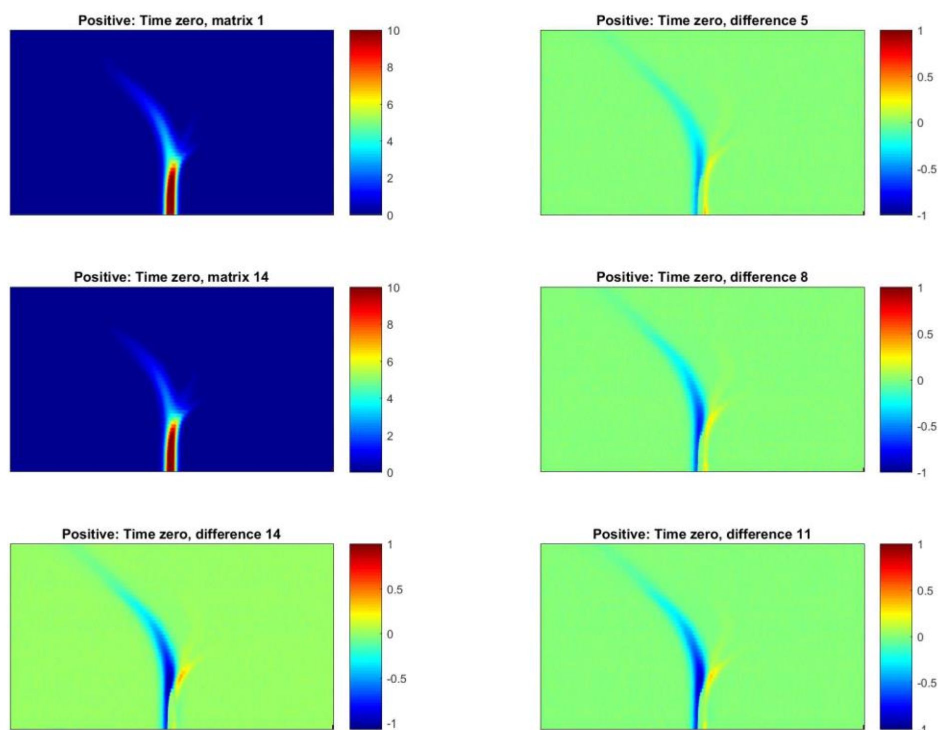


Figure 4.17: Positive FAIMS matrices: difference from baseline with matrix number: *Top left: First FAIMS spectrum at time zero. Middle left: Fourteenth FAIMS spectrum at time zero. Bottom left: Difference between fourteenth and first spectrum. Right hand panels show the difference for matrices 5, 8, and 11 from baseline, showing a consistent pattern, developing with time. All matrices averaged over all patients.*



4.5 Discussion

This developmental work has provided unique insight into the complexities of maintaining analytical consistency when detecting VOCs and also the effect of different sample storage procedures on sample analysis. Clear alterations in sample results with time left at room temperature were demonstrated. But it is not clear if this relates to a loss of VOCs and, most importantly, whether if there is a loss, the lost VOCs are crucial to disease detection.

It would seem perhaps logical, that urine samples would degrade at an accelerated rate at room temperature as opposed to when frozen as biological activity is higher with increasing temperature. The variation in the Lonestar spectra seem to increase monotonically with the age of the urine, with the exception of a plateau phase

between hours 12-48. The total ion count also appears to grow monotonically with age. Both these changes suggest an effect on the VOC capture, but more work is needed to establish which VOCs are affected and whether it affects the sensitivity of the Lonestar to detect specific disease. The increase in ion count is interesting: one might expect a decrease in total ions if it is assumed that VOCs are lost, but it depends where these ions are generated. It may be the result of sample degradation at room temperature. It is also possible that the presence of bacteria in the urine is causing biological breakdown of substances resulting in an increase in VOC's. However, in the absence of infection, urine should be sterile. There is often contamination on exit from the body, but it seems unlikely it would explain the pattern of consistent increased ion count and variability present in all samples. Increased variability with time supports the notion that a urine breakdown process is occurring, resulting in more small ions. It maybe that the plastic universal container may produce more VOCs whilst in contact with urine at room temperature, than would be observed while in frozen conditions.

When observing the effect of multiple samples runs though the Lonestar, in all samples there was consistently an increase in variation and reduced ion count with each matrix, except for a transient rise in ion count in the negative matrices. This may suggest that most information is gained from the first matrix, and subsequent runs may not add additional useful information in gleaning the VOC profile of urine in any particular disease. More work would need to be carried out to explore this finding before any changes to run numbers were made. In particular, a case-control study as an initial investigation, would establish whether key signals were lost after the first couple of runs. In the urine validation work documented in experiment one, an increase in ion count was found initially with a decrease at run four. However, it is important to note that the urine degradation samples in experiment four were run on a later version of Lonestar, this also had the addition of an auto sampler. Due to the differences in run speed and sampling technique between manual and auto sampling, it is likely results using differing sampling techniques cannot be directly compared.

Based on these preliminary results, ideally urine should be stored at -80^{0}C as soon as possible until further investigation into the nature of the lost VOCs is established and whether they are key to disease detection or not.

4.6 Conclusion and further work

VOC analysis using FAIMS is still in the experimental stage, but represents an exciting technology that has the potential to herald a new wave of non-invasive testing for a multitude of diseases. For FAIMS to be a real competitor as a diagnostic tool in the clinical setting it is important that minor changes in sample collection, storage and analysis technique or environment do not affect disease detection. The conflicting results seen between experiments one and four with regard to the effect of matrix/run number on ion count, also highlights the importance of consistency when testing samples as it may be that the auto sampling method cannot be compared to the manual method. These results provide new insight into the effect of sampling and storage on urinary analysis. However, these findings need to be corroborated in a case-control study to understand the impact on disease detection as changes to total ions and variability do not necessarily equate to an influence on the ability to detect a specific disease.

As a result of the experimental work described in this chapter, the following recommendations for samples analysis using the Lonestar FAIMS are:

1. Where possible, no changes should be made to any step of analysis during an experiment as even minor methodological changes have the potential to confound results.
2. Following sample collection, transfer to -80^{0}C should occur as soon as possible, but as a minimum within 12 hours of collection.
3. Samples should be stored at -80^{0}C for a maximum of 12 months prior to analysis.

CHAPTER 5:

The application of urinary volatile organic compound analysis in the detection of colorectal cancer

5.1 Introduction

The search for a non-invasive test that is both sensitive and specific in detecting and diagnosing CRC at an earlier stage, is ongoing. Applying the use of gas phase markers to diagnosis CRC is a rapidly expanding area. Over the past decade, many research groups have been focussed on the detection of VOCs in health and disease. Some disease-specific patterns have already been identified (200, 359). Yet there are currently no VOC-based tests used in the clinical setting (164, 229, 360).

Using urine for VOC detection has the advantage of being simple to collect, with high patient satisfaction. It is also easy to store and shows stability in the medium term (203). These factors make it an ideal focus for the development of a diagnostic test. The benefits of urine over other biological samples are discussed in Chapter Two.

Detection of VOCs in CRC using a variety of different mediums, has been investigated in small-scale studies (238, 240), but few have used urine analysis and none have focussed on the asymptomatic bowel cancer screening programme (BCSP) cohort. The BCSP uses a FOBT for CRC detection and has reduced mortality from CRC by increasing the proportion of patients being diagnosed at an earlier stage. However the full potential of bowel screening is limited by modest patient uptake (around 60%) and low specificity, which results in many unnecessary colonoscopies (117). The introduction of FIT as the first-line bowel screening test is anticipated within the next year and is associated with higher uptake and higher sensitivity, but there is still substantial room for improvement in both patient uptake and overall test accuracy. As VOC detection in the diagnosis of CRC in the asymptomatic population is currently unchartered territory, it is not clear whether a VOC based screening test would be feasible or effective.

In order to address these gaps in knowledge and carry out a comprehensive investigation of urinary VOC analysis in the detection of CRC, a series of three experiments were conducted. Firstly, in experiment one a large symptomatic patient group were included in order to investigate the detection of disease versus control, but also to establish whether different CRC groups could be classified based on

tumour stage. The second and third experiments concentrated on the asymptomatic BCSP population. Two different experiments were performed in the BCSP cohort because methodological improvements in sampling and analysis were utilised and, in addition, the first VOC analysis using GC-IMS was performed.

5.2 Hypothesis

The urinary VOC profile in CRC can be distinguished from the urinary VOC profile of benign bowel diseases.

5.3 Aims

1. To evaluate the use of VOC biomarkers to stage CRC and differentiate between cancer and advanced adenomas.
2. To establish whether urinary VOCs can form the basis of a new bowel cancer screening tool that is sensitive and specific, with high patient acceptability.

5.4 Research questions

1) a) Does CRC have a specific VOC biomarker 'fingerprint'?

   b) Can this be applied as a diagnostic test for CRC in the bowel cancer screening population?

   c) What is the sensitivity and specificity of this test in the screening population?

2 a) Can urinary VOC biomarkers differentiate between different stages of CRC?

   b) Can urinary VOC biomarkers differentiate between colorectal adenomas and CRC?

5.5 Ethical approval

Scientific and ethical approval was obtained from the Warwickshire Research & Development Department and Warwickshire Ethics Committee 09/H1211/38. Written informed consent was obtained from all patients who participated in the

study. Approval was granted by the Bowel Cancer Screening Programmed Research Committee.

5.6 Machine Learning

The analysis of data created through FAIMS and GC-IMS analysis employs machine learning methods. In brief, they involve the construction of computerised algorithms that can learn from and make predictions on the output data from the instrument. The algorithms build a prediction model from a training set of known cases and use this knowledge to predict output decisions (diagnoses) on unknown cases. The construction of these algorithms is described in more detail in Chapter Four section 4.2.3.

5.7 Experimental work

5.7.1 Experiment one: The application of urinary VOC analysis in the detection and differentiation of CRC by tumour stage in the symptomatic population: A retrospective case-control study.

Data analysis: August 2016

Study aim: To establish whether CRC can be correctly classified from control and whether the VOC profile is affected by cancer stage. Setting: UHCW. Inclusion criteria: Patients at UHCW diagnosed with colorectal adenocarcinoma of any stage between 2012-2014. Exclusion criteria: Any other histological cancer of the colorectum.

Method

Patient recruitment: The UHCW tissue bank database was used to identify all samples collected from patients with CRC in long-term storage at -80°C. Recruitment was conducted by the gastroenterology research nurses at UHCW between 2012-2014. Cancer cases were found by weekly review of the colorectal

multi-disciplinary team meeting at UHCW. Patients with a new diagnosis of CRC were contacted by phone to determine their interest in the study. Those who agreed, were approached at UHCW on the day of surgery prior to attending theatre. Patient characteristic data, co-morbidities and medication history were collected according to a pre-defined CRF and patients were consented. Control samples were collected between 2012-2014 from un-selected colorectal outpatient clinics in cases where no CRC, colonic adenomas or other gastrointestinal conditions were diagnosed and where a diagnosis of IBS was made**.** As per CRC samples, IBS samples were placed into long-term storage at -80 ^{0}C.

Sample collection and storage: 2x 20ml samples of urine were collected. Samples were immediately transferred to -20°C storage and then to -80 ^{o}C within 24 hours for long-term storage.

Sample analysis: All samples were analysed in august 2016 on the Lonestar instrument using an auto-sampler (Biomedical sensors laboratory, School of Engineering, University of Warwick). Please refer to Chapter Four for the detailed methodology**.** Outcome data including cancer stage was matched with VOC profile and compared to control sample VOC profile. Lonestar data was processed using method one, as described in Chapter Four.

Results

Study population: A total of 77 CRC samples and 109 IBS samples were analysed. Demographic details and major comorbidities are displayed in Table 5.1. The majority of cancers were found in the left colon (Table 5.2), Dukes C was the most common stage of presentation, seen in 36.4% of patients (Table 5.3).

Table 5.1. Study population characteristics for CRC and IBS groups

	CRC	IBS
% Male	60.0	27.5
Median age (mean)	69 (66.6)	44 (45.9)
Current smoker (%)	4 (5.2)	18 (16.6)
Median BMI	27.8	25.5
Median alcohol consumption (units/week)	6	0
Co-morbidity	**Number (%) affected**	
Hypertension	36 (46.8)	13 (12)
Diabetes	9 (11.7)	3 (2.7)
Irritable bowel syndrome	0	103 (100)
Inflammatory bowel disease	1 (1.13)	0

5. 2. Cancer distribution by site.

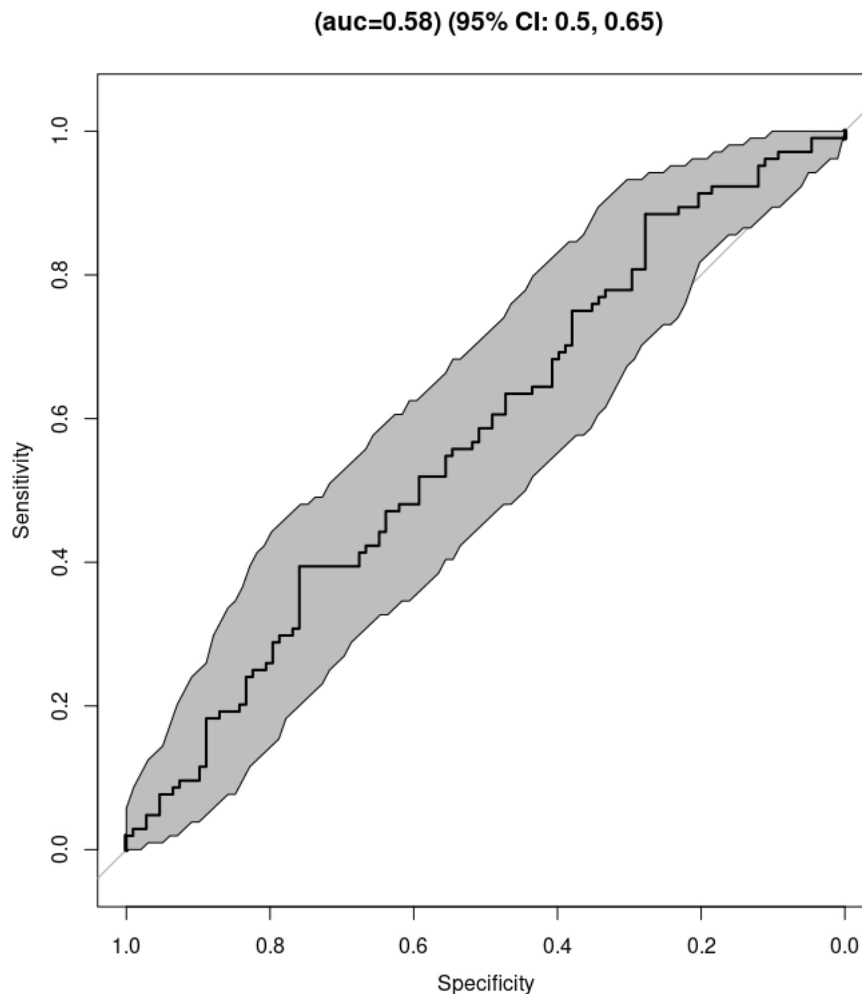
Cancer site	No. of patient (%)
Rectum	36 (46.8)
Sigmoid colon	8 (10.4)
Descending colon	2 (2.6)
Splenic flexure	3 (3.9)
Transverse colon	4 (5.2)
Hepatic flexure	7 (9.0)
Ascending colon	11 (14.3)
Caecum	6 (7.8)

Table 5.3. Stage of cancer according to Dukes criteria.

Dukes stage	No. of patients (%)
A	20 (26.0)
B	21 (27.2)
C	28 (36.4)
D	8 (10.4)

Cancer detection: When comparing the FAIMS matrices for cancer versus control, reclassification was found to be no better than random with AUC of 0.58, 95% CI 0.5-0.65 (figure 5.1). In view of the lack of separation demonstrated between the two groups no attempts were made to sub-analyse the data based on cancer stage.

Figure 5.1. ROC curve for CRC vs control in the symptomatic population. *The Middle line represents the sample point estimates, the outer lines represent 95% CI's.*



Discussion

This analysis of a large number of CRC samples versus control using the Lonestar instrument, demonstrated no ability to correctly classify CRC from the control. This means the VOC fingerprint of CRC could not be distinguished from control. Separation between disease and control has been demonstrated in past studies using FAIMS, and in particular in small-scale studies involving CRC, but these results did not replicate those findings.

There are a number of possibilities to consider when interpreting these results that may have influenced outcome. Firstly sample age: the samples had been in long term storage prior to sample analysis, this was due to unavoidable logistic and technical delays. The analysis was performed at the first available opportunity and both cancer and control samples were collected and stored over the same time period, so conditions prior to analysis were comparable. Although research by the FAMISHED group has suggested a decline in VOC number even at -80 ^{0} C (358). There is no data to confirm whether key VOCs are lost in this process, or whether it is simply a decrease in total VOCs, many of which may represent 'background noise'. These results support the argument that vital signalling information is lost as the samples ages. It is also important when considering a VOC based test in clinical practice, as it is likely samples would need an expiry date, after which results are invalid.

Secondly, these samples were analysed using an auto-sampling method. Theoretically there is no element of this sampling model that should affect the VOC detection, but samples are out of the fridge for longer than they would be via manual sampling (although on a cooling tray). Test experiments performed within the laboratory did not pick up any issues with the auto-sampling method, but this was the first analysis set involving large numbers of samples. Auto-sampling is a necessary step in the application of FAIMS in biological samples analysis, as manual testing is time consuming and cannot be feasibly carried out on large numbers. However, more work may need to be done to optimise sample processing using this method and to establish whether aspects of the auto-sampling process need refinement.

5.7.2 Experiment two: A prospective cohort study in the BCSP population, comparing urinary VOC profiles with diagnostic outcomes: (phase one).

Data analysis: February 2016

Study aim: To establish whether patients with cancer, adenoma or normal colons in the BCSP population could be correctly classified based on their VOC profile using the Lonestar. Setting: UHCW. Inclusion criteria: Any patient attending the BCSP clinic at UHCW following a positive FOBT. Exclusion criteria: Patients who declined investigation as per BCSP protocol, as no outcome data would be available. Patients on renal dialysis as the effect of renal replacement therapy on VOC profile is unknown.

Method

Patient recruitment: All patients with a positive FOBT attending an appointment with a bowel cancer screening nurse at UCHW were eligible for inclusion. Eligible patients were highlighted using the trust clinic list database. Both the local UHCW bowel cancer screening team and the bowel cancer screening hub in Rugby, England, were approached in order to facilitate recruitment. Recruitment was conducted at the UHCW screening clinic between April 2015 and February 2016. All patients who agreed to further investigations were recruited.

Sample collection and storage: As per experiment one.

Sample analysis: Detailed methodology has not been repeated here. Manual sampling using the Lonestar instrument was performed as per the method outlined in Chapter Four. Statistical analysis was performed using method one from Chapter Four. Outcome data was collected either from the colonoscopy or CTC report. Any cancer detected was cross-checked with the histology and staging was determined via detail from the MDT meeting and oncological correspondence. Any detected polyps were cross-checked with the histology report and included only if adenoma was confirmed. Hyperplastic polyps were excluded from analysis.

Results

127 patients were recruited, 125 samples were available for analysis. 57% were male, and the median age was 68. Average BMI was 28.73. 120 (96%) stated that their diet was normal, 5 (4%) were vegetarian, 1 (0.80%) ate no red meat and 1 (0.80%) ate no beef. Average weekly exercise duration was 3.08 hours but the median was 0 hours. Average weekly alcohol intake was 11.25 units; the median was 4. 13 (10.4%) were taking probiotics. Table 5.4 details major co-morbidities and those on anti-platelets/anti-coagulants or Non-Steroidal Anti Inflammatory Drugs (NSAIDS).

Table 5.4. Study population characteristics and medications for BCSP patients

Condition	Number (%) affected
Hypertension	60 (48%)
Diabetes	16 (12.8%) all type 2
Irritable bowel syndrome	7 (5.6%)
Inflammatory bowel disease	2 (1.6%) both ulcerative colitis
Drug	**No. of patients**
Aspirin	12
NSAIDS	8
Warfarin	7
Other antiplatelets	6

Table 5.5 lists the patient diagnosis. Where applicable, the cancer site and adenoma risk category are stated. The majority of cancers were in the left colon which is consistent with the national trend.

Table 5.5: Diagnostic outcomes in the BCSP cohort *(A total of 12 cancers as one patient had synchronous tumours in the caecum and sigmoid)*

Diagnosis	**No. patients (%)**	
CRC	Total	11 (8.8)
	Rectum	5 (41.6
	Sigmoid	3 (25)
	Descending	0
	Transverse	1 (8.4)
	Ascending	1 (8.4)
	Caecum	2 (16.6)
Adenomas	Total	55 (44)
	High risk	10 (8)
	Intermediate risk	22 (17.6)
	Low risk	23 (18.4)
Diverticular disease	12 (9.6)	
Haemorrhoids	5 (4)	
Normal	27 (21.6)	
Other	15 (12)	

Sample groups were paired using five different combinations based on disease outcome: numbers in brackets denote sample size. Test accuracy in correctly classifying cases into the each group is displayed as AUC below. This data is also shown in Table 5.6, together with the sensitivity and specificity values and corresponding 95% CIs.

a) Adenoma (55) vs no Adenoma (70) – the adenoma group included all those with adenomas regardless of risk category. No adenoma group, included all other samples without the presence of adenoma. AUC was 0.554 with 95% CI 0.451-0.657.

b) Adenoma (55) vs no disease (27) - no disease group included all those with an outcome of normal. AUC was 0.548 with 95% CI 0.44-0.655.

c) <u>Cancer (11) vs no disease (27)</u> – here all samples from patients with a diagnosis of CRC were compared with samples with an outcome of normal. AUC was 0.660 with 95% CI 0.46-0.87.

d) <u>Cancer (11) vs no cancer (114)</u>- patients with cancer were compared with all remaining samples that had any other diagnosis apart from cancer. AUC was 0.636 with 95% CI 0.497-0.776.

e) <u>Disease (98) vs no disease (27)</u> - patients with any diagnosis including cancer were compared with samples where no diagnosis was found and the reported outcome was normal. AUC was 0.571 with 95% CI 0.468-0.674.

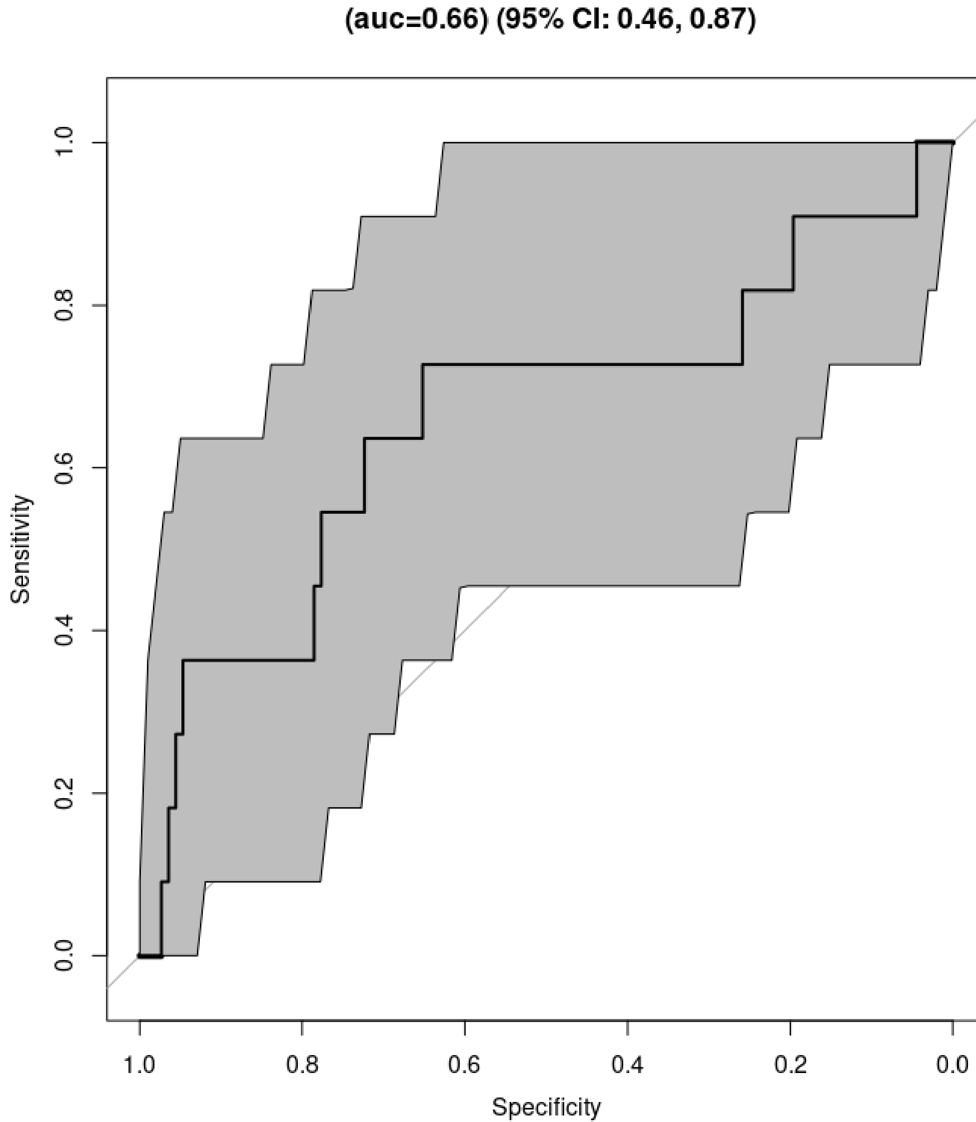
Table 5.6: Classification of BCSP cohort by outcome: *AUC, sensitivity and specificity with 95% CI.*

Groups	AUC (CI)	Sensitivity (C.I)	Specificity (C.I)
a) Adenoma vs no Adenoma	0.554 (0.451-0.657)	0.582 (0.421-0.531)	0.574 (0.494-0.593)
b) Adenoma vs no disease	0.548 (0.44-0.655)	0.545 (0.421-0531)	0.552 (0.509-0.616)
c) Cancer vs no disease	0.660 (0.46-0.87)	0.636 (0.459-0.73)	0.655 (0.601-0.703)
d) Cancer vs no cancer	0.636 (0.497-0.776)	0.636 (0.571-0.824)	0.598 (0.359-0.442)
e) Disease vs no disease	0.571 (0.468-0.674)	0.538 (0.493-0.594)	0.534 (0.518-0.624)

Using this classifier test, no significant separation between any of the five groups based on VOC profile was demonstrated. The groups c) Cancer vs no disease classifiers and d) Cancer vs no cancer, displayed 95% CI's that just include 0.5. This was due to the small sample size and suggests statistical discriminative power is poor between groups (table 5.6). Figure 5.2 shows graphically the poor

classification of cancer versus no disease curve. Larger numbers would be required to ascertain if there was any significant difference between the two groups. The sensitivity and specificity data was consistent between groups with the exception of c) cancer vs no disease where a slightly higher sensitivity of 0.636 with 95% CI 0.459-0.73 was seen. None were statistically significant. Although this falls well below the values expected for a satisfactory test, it suggests that a larger sample size would be required to investigate this further. Adenoma compared with other groups was also poorly classified, therefore further sub-analysis based on adenoma risk category was not performed.

***Figure 5.2. ROC curve for CRC vs no disease in the BCSP cohort** using sparse logistic regression. The bold central line represents the sample point estimate and the outer lines represent the wide 95% CI's, suggesting poor separation of groups.*



Discussion

The results of this investigation into the analysis of VOCs in the FOBT positive BCSP population did not show any satisfactory degree of separation between cancer and any other group or adenoma vs any other group. Sensitivity and specificity results were comparable. Therefore it was not possible to correctly classify patients into the correct group based on their diagnosis. Although the total group size was

large in comparison to previous VOC-based studies, the cancer cohort was small and reflects the low incidence of cancer in the screening population (approximately 8%) (117), In light of these results, recruitment was continued to improve sample size and subsequently perform re-analysis. These results are described in experiment three.

5.7.3 Experiment three: A prospective cohort study in the bowel cancer screening population comparing urinary VOC profiles with diagnostic outcomes (phase two).

Data analysis: October 2017

As with phase one of this study, the experiment included patients recruited via the BCSP at UHCW. Therefore the aims and research questions remain the same. This experiment was repeated for the following reasons:

1. Recruitment was extended from April 2015 to November 2016, resulting in a larger sample size.
2. Lonestar development: since experiment one the Lonestar instrument had a software update and system clean.
3. Data analytical techniques also improved, allowing more accurate classification and extraction of key data features, resulting in higher performing prediction models.
4. Access to GC-IMS analysis became available, allowing evaluation of a second instrument for VOC detection.

Method

Sample collection and storage: this remained unchanged and is outlined in the earlier section. Sample analysis: manual analysis took place in October 2017 using the Lonestar v3.1. In addition, GC-IMS analysis was conducted on a sub-set of patients in the same laboratory. Detailed methodology plus data analysis methods for GC-IMS are found in Chapter Four. Statistical analysis: method three was used. As cancer numbers were small, a balancing technique was applied to the data in order to fairly match the non-CRC samples with the same number of CRC samples. Rather than randomly selecting a set number of controls to compare to the CRC

group, balancing involved a Synthetic Minority Over-Sampling Technique (SMOTE), (361) where artificially generated points are plotted to represent the control group as a whole and is used to provide a more fair representation. Patients were grouped into six categories for analysis. In order to extrapolate the most clinically relevant comparisons, study numbers are denoted in brackets.

a) Cancer (12) vs Normal (12): here the data has been balanced by matching the number of normal cases to that of the control, as described above.

b) Cancer +all adenomas (93) vs normal (37)

c) Cancer + high risk adenomas (30) vs normal (37)

d) Cancer + high risk adenomas (30) vs other (70)

e) Cancer + all adenomas (93) vs other (70)

f) Non- cancer (113) vs normal (37): ‘non-cancer’ describes all polyps plus other diagnoses except normal and cancer. ‘Other’ includes all other diagnoses plus normal

Results

A total of 201 samples were collected, however, following the use of samples for experiment one in February 2016, 171 were available for phase two. A further 8 patients were excluded because there was no outcome result, therefore a total of 163 samples were analysed. 93 (57%) were male, average age was 66.57 median age was 67 years.12 (7.4%) were current smokers. 41 (25.4%) were ex-smokers and 109 (67.2%) had never smoked (there was no data on one patient). Table 5.7 lists the diagnosis for the group analysed.

Table 5.7. ***Diagnostic outcomes for BCSP cohort (phase two)*** *and distribution of cancer by site (total of 13 cancer sites as one patient had synchronous tumours.)*

Diagnosis	Number (%)	
Cancer	Total	12 (7.6)
	Rectum	4 (2.4)
	Sigmoid	4 (2.4)
	Descending colon	0
	Transverse colon	1 (0.58)
	Ascending colon	2 (1.17)
	Caecum	2 (1.17)
Adenoma	Total	80 (49.1)
	High	17(10.5)
	Intermediate	36 (21.1)
	Low	27 (17.5)
Diverticular disease		14 (8.2)
Normal		37 (19.3)
Haemorrhoids		5 (2.9)
Other		14^ (8.2)
Excluded		8* (4.7)

^7 IBD, 2 rectal telangiectasia, 1 rectal ulcer, 1 radiation proctitis, 1 inflammatory pseudopolyp, 1 non-specific sigmoid inflammation, 1 ischaemic sigmoid stricture
*1 not fit enough for investigations, 7 declined investigations

For each group, the results for the classification model that demonstrated the best separation are displayed in Table 5.8.

Group a) Cancer vs normal demonstrated the highest degree of separation with AUC 0.98 95% CI 0.93-1 (figure 5.3) with 12 patients in each group. Sensitivity and specificity were also high: 1 95% CI 0.74-1 and 0.92 95% 0.62-1 respectively p=0.000002.

In groups b) - e) cancer was grouped with adenomas and showed only modest AUC and corresponding sensitivity and specificity without any statistical significance. The highest separation of the adenoma groups was seen in b) cancer + all adenomas

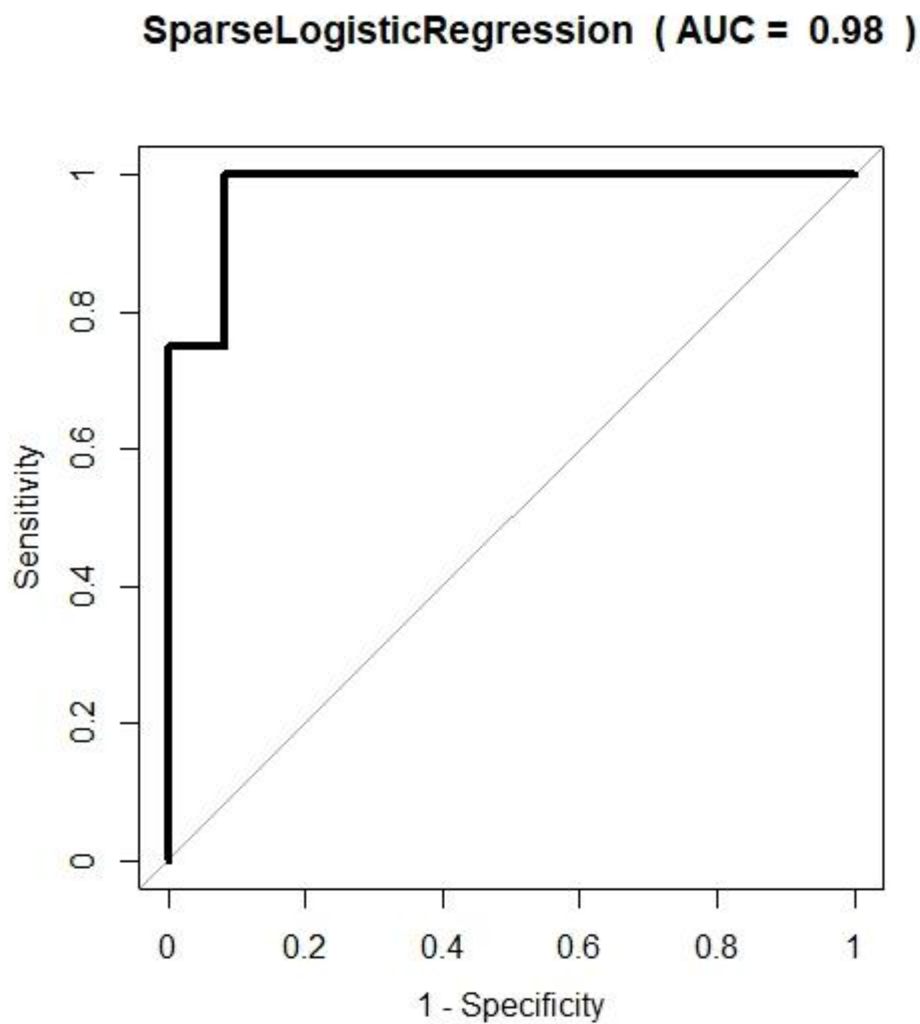
vs normal; sensitivity was low at 0.48 95% CI 0.38-0.59 but specificity was high at 0.89 95% CI 0.75-0.97 *p=0.9937*. In f) when cancer was excluded, separation was low with sensitivity 56%, 95% CI 0.46-0.65.

A further set of analyses were carried out to investigate the classification of the adenoma groups in more depth and to compare cancer with three categories of adenoma: a) High risk, b) Intermediate risk and c) Low risk, according to the BSG guidelines (table 5.9). Numbers were small, but high sensitivity was demonstrated when each adenoma group was compared with cancer. The most accurate overall classification was seen in cancer vs high risk adenoma with a sensitivity of 0.83 95% CI 0.52-0.98 p=0.0006.

Table 5.8. Classification of BCSP cohort (phase two) by outcome. *Using the best feature classifier with corresponding 95 % CI. Numbers in brackets in group section denote sample number.*

Group	Feature Classifier	AUC	Sensitivity	Specificity	PPV	NPV
a) Cancer (12) vs normal (12)	Sparse Logistic Regression	0.98 (0.93-1)	1 (0.74-1)	0.92 (0.62-1)	0.92	1
b) Cancer + all adenomas (93) vs normal (37)	Sparse Logistic Regression	0.64 (0.54-0.74)	0.48 (0.38-0.59)	0.89 (0.75-0.97)	0.92	0.41
c) Cancer + high risk adenomas (30) vs normal (37)	Gaussian Process	0.62 (0.48-0.76)	0.57 (0.37-0.75)	0.68 (0.5-0.82)	0.59	0.66
d) Cancer +high risk adenomas (30) vs other (70)	Sparse Logistic Regression	0.6 (0.47-0.73)	0.47 (0.28-0.66)	0.80 (0.68-0.89)	0.52	0.76
e) Cancer + all adenomas (93)vs other (70)	Sparse Logistic Regression	0.56 (0.47-0.65)	0.91 (0.84-0.96)	0.25 (0.15-0.38)	0.64	0.67
f) Non cancer (113) vs normal (37)	Gaussian Process	0.61 (0.51-0.71)	0.56 (0.46-0.65)	0.68 (0.5-0.82)	0.83	0.35

Figure 5.3. ROC curve for classification of CRC vs normal in BCSP cohort *(balanced) using SLR classifier. The bold line on the left of the graph represents the point estimate for the samples analysed (the shape of the line is due to the low number of samples) and indicates a high degree of accuracy for separation of the two groups.*



.

Table 5.9. Classification of BCSP cohort: CRC vs adenoma *(hr-high risk ir- intermediate risk lr- low risk) using the best feature classifier with corresponding 95% CI*

Group	Feature Classifier	AUC	Sensitivity	Specificity	PPV	NPV
a) Cancer (12) vs Adenoma (7) (hr)	Random Forest	0.92 (0.77-1)	0.83 (0.52-0.98)	1 (0.59-1)	1	0.78
b)Cancer (12) vs Adenoma (12) (ir)	Random Forest	0.84 (0.67-1)	0.83 (0.52-0.98)	0.75 (0.43-0.95)	0.77	0.82
c)Cancer (12) vs Adenoma (12) (lr)	Gaussion Process	0.83 (0.66-1)	0.75 (0.43-0.95)	0.92 (0.62-1)	0.90	0.79

GC-IMS analysis

A subset of 109 patient samples were analysed through the Silox GC-IMS instrument. Five comparator groups were devised according to outcome: a) Cancer vs normal, b) Cancer+ high risk adenoma vs normal, c) Cancer vs other, d) Cancer+ all adenomas vs other e) All adenomas vs normal. All five feature classifiers were used, the result with the best fit is displayed in Table 5.10.

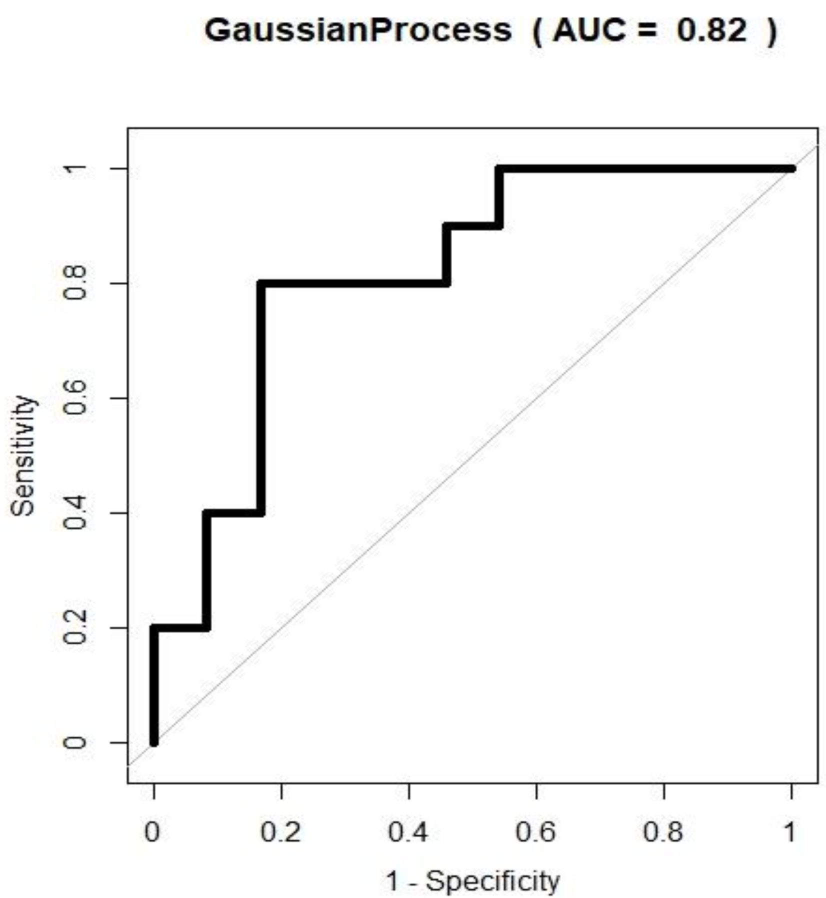
.

Table 5.10. Classification of BCSP cohort using GC-IMS *using best feature classifier with corresponding 95 % CI. Sample numbers are denoted in brackets*

Group	Feature Classifier	AUC	Sensitivity	Specificity	PPV	NPV
a) Cancer (10) vs normal (24)	Gaussian Process	0.82 (0.67-0.97)	0.80 (0.44-0.97)	0.83 (0.63-0.95)	0.67	0.91
b) Cancer + high risk adenomas (23) vs normal(24)	Gaussian Process	0.53 (0.36-0.70)	0.48 (0.27-0.69)	0.67 (0.45-0.84)	0.58	0.57
c) Cancer (10) vs other (20)	Support Vector Machine	0.77 (0.60-0.94)	1 (0.66-1)	0.57 (0.34-0.78)	0.5	1
d) Cancer + all adenomas (65) vs other (42)	Support Vector Machine	0.61 (0.49-0.72)	0.71 (0.58-0.81)	0.55 (0.39-0.70)	0.71	0.55
e) All adenomas (55) vs normal (24)	Gaussian Process	0.61 (0.47-0.75)	0.58 (0.44-0.71)	0.62 (0.41-0.81)	0.78	0.39

As with the analysis using FAIMS, when comparing cancer versus normal (group a) there was a high degree of separation with a sensitivity of 80%, 95% CI 0.44-0.97 and specificity of 83%, 95% CI 0.63-0.95 but with a *p value* of 0.9986, the corresponding ROC curve is seen in figure 5.4. Cancer vs other also had a high sensitivity of 100% 95% CI 0.66-1 but specificity dropped to 57%, 95% CI 0.34-0.78. When cancer samples were grouped with adenomas and compared with 'other' groups the sensitivity dropped to a modest level of 71%, 95% CI 0.58-0.81. Adenomas vs normal showed a low level of separation, with a sensitivity of only 58% 95% CI 0.44-0.71.

Figure 5.4: ROC curve for classification of CR-C vs normal using GC-IMS.
(GP classifier). The bold line is the point estimate for the samples analysed. The sample line is to the left of the AUC=0.5 line which is inserted for comparison. This emphasises the high test accuracy found in separating these two groups.



The GC-IMS produces thousands of data points following analysis of a sample, this data can be presented as a heatmap which can represent the raw data graphically. This is shown in figures 5a and 5b. The drift time on the X axis comes from GC and the retention time on the Y axis comes from IMS. Feature selection is a step in the data analysis algorithm outlined in Chapter Four. The purpose of feature selection is to remove data that is outside of a defined threshold for which important information is contained. It removes data that doesn't contain useful information and may represent background data due to environmental changes. The change in heatmap before and after feature selection demonstrates how the algorithm used to manage the raw data 'cleans' up the points, focussing in on those of interest.

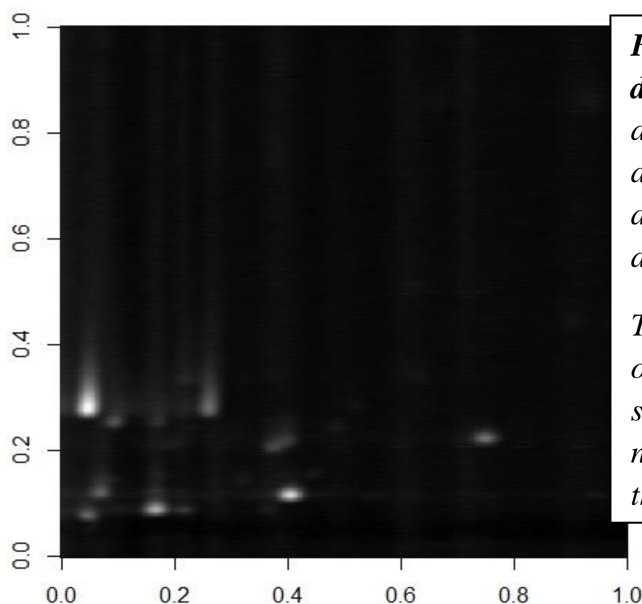


Figure 5.5a: GC-IMS raw output data before feature selection *applied. X axis – drift time (M/s), Y axis – retention time (M/s.) White areas represent clusters of ionic activity.*

These figures demonstrate the actual output data from running a urine sample before conversion to numerical values and represents thousands of data points.

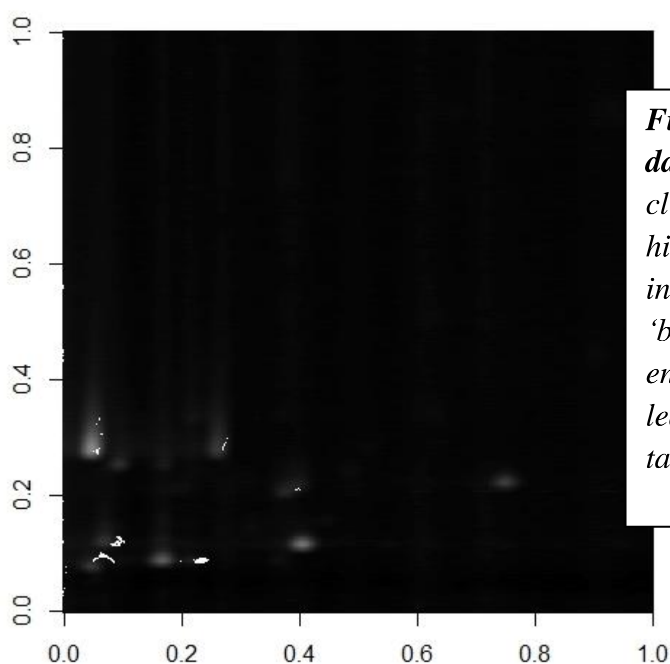


Figure 5.5b: GC-IMS raw output data after feature selection*. White clusters are more well defined, highlighting specific points of interest. Data representing 'background noise' due to environmental factors is removed, leaving refined and more well targeted data points.*

5.8 Discussion

Overall this experimental work, which examined the use of urinary VOC markers as a tool to detect CRC, produced varied results. This is, in part, a reflection of the evolving understanding of VOC analysis and the ongoing improvement of sample and data analysis.

FAIMS analysis

Results using the most up to date methodology for FAIMS analysis (experiment three) in the FOBT positive BCSP population showed:

- Classification of CRC from normal was high with AUC 0.98, sensitivity of 1 and specificity of 0.92
- Classification of CRC from low, intermediate and high risk adenomas was high with test accuracy ranging from 0.83-0.92.
- When CRC was grouped with adenomas and attempts made to classify this combined group with other or normal groups the separation was low with test accuracy ranging from 0.56-0.64.

GC-IMS analysis

Results of the first study using GC-IMS to detect CRC showed:

- Classification of CRC from normal was high with AUC 0.82, sensitivity of 0.80 and specificity of 0.83.
- Classification of CRC from other disease groups demonstrated an AUC 0.77, sensitivity of 1 and a low specificity of 0.57.
- When CRC was grouped with adenomas test accuracy dropped to 0.53-0.61.
- Classification of adenoma from normal was low with AUC 0.61, sensitivity 0.58 and specificity 0.62.

The results from both FAIMS and GC-IMS show consistency. The separation of CRC from normal cases was high, yet when CRC cases were grouped with adenomas the accuracy dropped significantly. This suggests that CRC has a VOC profile that distinguishes it from other GI pathologies, and therefore when this

profile is combined with that of other disease groups the VOC signal is not distinct enough to be correctly classified from normal groups. The observation that CRC and adenomas have distinct profiles, was further supported by the analysis using FAIMS to separate cancer from adenomas. In this case the separation was high, which suggests that it is possible to separate malignant from pre-malignant disease based on urinary VOC profile. Specificity when comparing CRC and adenoma groups was particularly high. This set of results warrant further exploration to ascertain whether the separation seen is simply because the adenoma group represent another non-cancer group or whether there are changes specific to the VOC profile of colonic adenomas that make it more distinct from CRC than other groups.

CRC-specific VOCs are thought to occur via genetic and protein changes that cause peroxidation of the cell membrane. In addition, there is an increase in reactive oxygen species within the cancer cell. Finally, alterations in the microbiome have a direct effect on the VOCs. As discussed in Chapter Two, there has already been work carried out by other research groups, to isolate the chemicals that may be specific to CRC cases. Currently, there is little consensus as to which hold most importance. Hydrogen Sulphide has also been implicated in CRC by mediating DNA damage (235), but this has also been associated with colonic adenoma formation (362). It is well established that colonic adenomas are linked to microbial dysbiosis (363), with higher proportions of *Proteobacteria* and less *Bacteroidetes* in those with adenomas compared with controls. This suggests changes in the microbial environment may be a factor in adenoma development. In established CRC, changes in microbes found at the tumour site have been observed compared with normal tissue (364, 365) such as increased *Fusobacterium,* making it plausible that dysbiosis could be the cause of adenomas but a consequence of CRC. If dysbiosis is the key driver of altered VOC patterns, this would help to explain the difference in VOC signals between adenomas and CRC seen in this work.

The experiments in this chapter focussed not only on cancer, but on the detection of colonic adenomas versus non-cancer groups. This is the first study to examine adenoma detection using the urinary VOC profile. The adenoma group showed poor

separation from normal cases. Few groups have published data on colonic adenomas. De Meij et al (166) carried out faecal analysis and found sensitivities of 62% when advanced adenoma was compared with control. However, unlike the results seen here, when adenomas and cancer cases were grouped together against control the sensitivity improved to 85%. Amal et al used breath testing to differentiate advanced adenomas from controls with sensitivity of 100% which performed better than cancer vs control where sensitivity was 85% (239). This suggests current understanding is conflicting and there is more work needed to establish the mechanism of VOC profile change with colonic adenomas and other GI diseases.

The analytical technology associated with VOC detection is improving constantly. Refinement in data software and ion capture, as seen in experiment three, is likely to explain, at least partially the inconsistent results seen in the three experiments presented here. This reinforces the importance of ongoing modifications of the analytical instruments used for VOC capture. Machine learning models that utilise data analysis algorithms of the VOC output must be able to transform raw data into points of interest, whilst excluding excess information that does not reflect the disease state. This complex statistical processing is constantly improving, and the transition to more sophisticated data analysis is outlined in Chapter Four and has been demonstrated in this chapter. There are many positive developments in sampling and data analysis, yet the variables affecting VOCs are complex, and to a certain extent, as yet unknown.

The benefits of using FAIMS as the detection method of choice, have been described in detail in Chapter Four, but the inconsistent results between experiments reflect the many factors that alter VOC detection. These include temperature, background VOCs and storage factors and as yet have not been fully controlled for. The level of repeatability required for equipping the clinical setting with this technology has not yet been demonstrated. A major drawback of this technology, is the lack of ability to identify chemical compounds using FAIMS. Chemical identification would help to focus VOC profiling more intently and

improve understanding of the pathology of CRC. Unlike FAIMS, GC-MS has the ability to identify the chemical profile of a biological sample.

It is widely accepted that GC-MS is the gold standard in VOC detection (176, 182). In contrast to the other modes of analysis described in this thesis that rely on a pattern recognitions, GC-MS has the ability to detect aberrant VOCs that may serve as diagnostic biomarkers of individual disease. However, the benefits are matched by disadvantages, these include size, running cost and operation requiring highly trained staff (356). Therefore, GC-MS cannot be referred to as the gold standard for clinical practice. In view of the difficulties with GC-MS, the multitude of technologies being used by different research groups for analysis will continue, until a consensus is reached on which are most suitable. GC-IMS is a relatively new instrument, but it has the ability to provide chemical information and it is portable and easy to use for urine analysis. This means it has great potential with encouraging results presented in this chapter. But progress is hampered somewhat due to a lack of operators with experience in biological sampling. Therefore its adaptation in the clinical field is still in its infancy, a chemical registry that allows identification of the chemical make-up of biological samples has not yet been established. The panacea for VOC analysis is still awaited.

Future research

Important considerations in the storage and sampling conditions of urine specimens have been raised in this chapter and these considerations apply to FAIMS, GC-IMS and any other VOC detection modality. As examined in Chapter Four, sample age is important when capturing VOCs. The varying age of the samples analysed is likely to have had an impact on the results. This applies to experiment one particularly as some samples were collected in 2012. The knowledge and understanding the optimal conditions for storage prior to analysis is still in the early phases and more work focussed solely on this would be welcomed. However, VOC mediated disease detection is going to be a real contender in the clinical environment, the VOCs of interest will need to be stable enough to ensure an element of variability in terms of storage procedure.

It is important to highlight that the bowel screening patient group recruited for experiments two and three were those that had a positive FOBT. The current guaiac FOBT has poor selectivity for CRC, therefore it would be interesting to repeat this experiment once the FIT is introduced as the current screening tool, because it has both superior sensitivity and specificity compared to the current FOBT. Also key when comparing a new screening tool with FOBT is to examine the FOBT negative patients too, however access to this group is difficult as it requires permission from the BCSP research committee and recruitment of patients in the community rather than a hospital setting.

Combining VOC detection with other biochemical tests such as FIT as a two-stage test in the symptomatic population (392) is a newly proposed area of interest and holds potential in the field of CRC diagnosis. Work in this field is expected in the next twelve months and heralds new wave of possibility in the use of VOC detection in disease.

Limitations

A major limitation with these experiments is the small sample sizes for the CRC groups, therefore only tentative conclusions can be made. In isolation, VOC-based diagnostic testing in CRC has not yet demonstrated consistency or accuracy on a large scale.

Recommendations

VOC based diagnostics holds huge potential with multiple advantages for the patients, but there is still extensive work to be done before this can become functional in today's NHS. It is a rapidly developing area and further large scale, studies are taking place in VOC breath analysis.

The following recommendations are made based on this work

An international collaborative approach would be welcomed to formulate consensus statements on sampling technique and instruments used in VOC detection. Also of vital importance is the development of a GC-MS based registry of chemicals isolated from CRC samples. This would contribute to a pattern-

recognition library for all diseases investigated for using VOC detection and would allow more focused analysis by narrowing the chemical window of interest. It would also contribute to the understanding of the pathophysiology of disease development, as investigation into the role of highly abundant chemicals could take place. Alongside these measures for improvement, it is also important to engage with manufacturers of the analytical instruments in order to motivate their focus on specific designs for human disease testing, rather than researchers relying on adapting industry-specific devices for biological samples

Research questions with answers

1 a) Does colorectal cancer have a specific VOC biomarker 'fingerprint'?

   The results from experiment three show the ability to correctly classify CRC samples from normal or 'other pathologies' with high sensitivity and modest specificity using FAIMS and GC-IMS. This suggests a chemical fingerprint that is unique to CRC. However the low study numbers must be taken into consideration.

   b) Can this be applied as a diagnostic test for colorectal cancer in the bowel cancer screening population?

   Recruitment for urine sampling was >90% suggesting a clear advantage of urine screening over faecal screening, but significant developments in methodology and analysis using VOC detection, are required before a suitable test is available in the clinical setting.

   c) What is the sensitivity and specificity of this test in the screening population?

   These results show a sensitivity between 80-100% and specificity 57-92% in FOBT positive patients. In comparison the current guaiac FOBT has a sensitivity of around 50% and specificity of 75%, therefore the urinary VOC test holds promise in this small group, but requires validation on a larger scale.

2 a) Can urinary VOC biomarkers differentiate between different stages of colorectal cancer?

This could not be established from the experiments performed in this chapter.

b) Can urinary VOC biomarkers differentiate between colorectal adenomas and CRC?

Analysis from this work indicates a high degree of separation using FAIMS VOC analysis for cancer vs adenomas with the largest separation seen between cancer and high risk adenomas (Sensitivity 0.83 95% CI 0.52-0.98, P=0.0006). This is an important finding that warrants further investigation and is also supported by the low sensitivities seen when cancer and adenomas are grouped together, suggesting they have substantially different VOC profiles.

5.9 Conclusion

The series of experiments described in this chapter investigates the feasibility of detecting CRC and adenomas using FAIMS for urinary VOC detection. This has not been replicated by any other research group in such high numbers, or in the BCSP population. In addition it is one of the first studies to report on the detection of colorectal diseases using GC-IMS. These results indicate that CRC can be correctly classified from control and adenomas using FAIMS and GC-IMS, but the classification of adenomas from control was poor. This approach to disease detection faces multiple challenges, that reflect the complexity of human disease and does not currently have a role in clinical diagnostics. There is a need for better encoding of features, future identification of the VOC composite molecules would be a huge step forward in achieving this. Collaborative research in the field of VOC diagnostics would allow investigation in more depth and with larger numbers than have currently been studied. This would more substantially address the plethora or variables that have an effect on this method of disease detection.

CHAPTER 6:

Investigating the urinary proteome in the detection of colorectal cancer

6.1 Introduction

The experimental work in this chapter demonstrates the first application of urinary peptide markers, to form a novel diagnostic test for CRC detection. It also evaluates the role of the proteome in the development of CRC.

Cancer can cause protein specific changes both local to the site of the malignancy and also systemically. Normally, daily urinary protein excretion is only <150mg/day (366) due to the process of reabsorption in the renal proximal tubules, but when cancer is present it has been hypothesised that the urinary proteome changes, and cancer-specific alterations may be detected as a marker of the presence of disease. Urinary proteomics is increasingly being used to identify new biomarkers of disease, but also to improve the understanding of the pathogenesis of disease. Urinary protein profiling has already been investigated for cancers of the urinary tract and pancreas (249, 367), but has not been explored in gastrointestinal cancers.

The heterogeneity between different colonic cancers has been acknowledged at the genomic and proteomic level, but also between tumour subtypes (368). Proteomic research has the potential to further explain the mechanisms of progression and prognosis of CRC that, as yet, have evaded researchers. Through focus on the proteome it may be possible to isolate key alterations in peptide sequencing that are specific to CRC. This could be achieved by identifying small numbers of disease-specific biomarkers from thousands of polypeptides using urine for analysis. Previously in this thesis, an outline of the benefits of urine as a source for disease detection or screening compared with other mediums, has been discussed. In addition to the advantages already mentioned, it has been postulated that the urinary proteome is less complex than other bodily fluids, such as blood, with a lower dynamic range (369), potentially making analysis easier.

The experiments performed in this chapter are multi-stage, with the aim of evaluating carcinoma-specific changes to the protein profile, in both the urine and colonic tissue of a patient with CRC. Firstly, a new technology in the medical diagnostic field, called capillary electrophoresis coupled with mass spectrometry

(CE-MS) was used. This is a highly sensitive method of identifying urinary peptides and their corresponding amino-acid chains. But it has rarely been applied to the field of urinary diagnostics before. Secondly, the marker peptides found in the urine were sequenced and linked to a target list of proteins. In the third stage of analysis, immunohistochemistry (IHC) was used to establish whether key proteins can be linked to the site of cancer by using an immunoassay applied to the colonic cancer tissue. The possible role of these proteins in the pathology of CRC was discussed.

6.2 Aims

1. To test the development of a urinary peptide diagnostic assay using CE-MS for the detection of CRC.
2. To identify proteins with a role in CRC, based on sequencing of the urinary peptides.
3. To establish if key proteins are present within the CRC tissue and discuss their role in the pathology of the disease.

6.3 Research questions

1. Do patients with CRC have a different urinary peptidome to controls?
2. What are the key proteins coded for by the peptide sequences detected in the urine of patients with CRC?
3. What is the potential role of these proteins at the cancer site?

6.4 Ethical approval

Scientific and ethical approval was obtained from the Warwickshire Research & Development Department and Warwickshire Ethics Committee 09/H1211/38. Written informed consent was obtained from all patients who participated in the study.

6.5 Experimental work

6.5.1 Identification of urinary peptide markers for CRC by capillary electrophoresis coupled to mass spectrometry.

Setting: Recruitment took place at UHCW, analysis was undertaken by Mosaiques-diagnostics in Hannover, Germany. Inclusion criteria: Colorectal adenocarcinoma and colonic adenoma cases that were confirmed histologically. Exclusion criteria: Cases where colonoscopy was not performed, therefore outcome unknown.

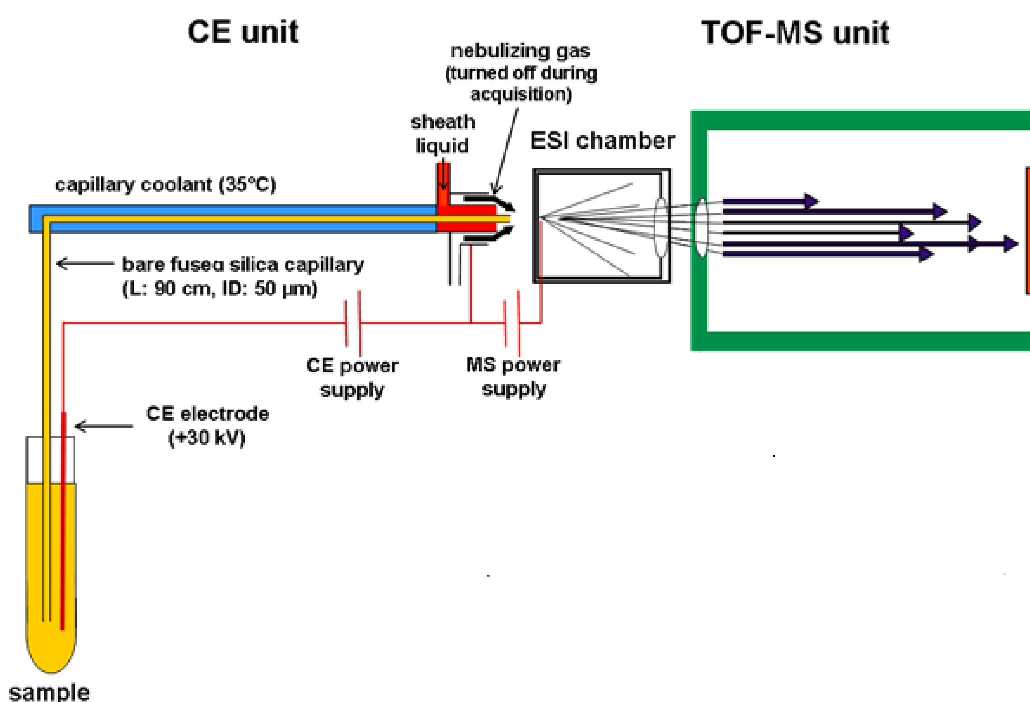
Method

Sample collection: During 2017, the FAMISHED database at UHCW was examined to identify 12 CRC cases, these were age-matched with 6 colonic adenoma cases and 6 normal controls. Sample collection and storage followed the method described in Chapter Four. Samples remained frozen using dry ice, whilst in transit to the study collaborators in Hannover, Germany (Mosaiques-diagnostics) where sample analysis took place in September 2017.

CE-MS

CE-MS was introduced in Chapter Two. It is a two-phase technique that separates molecules based on their mass-to-charge ratio and demonstrates several advantages for peptide detection. Schematics of the sample preparation and CE-MS process are shown in figure 6.1 below. CE-MS analysis methodology has been well described by previous research groups (247, 370) and has been replicated in this experimental work. CE-MS is highly sensitive to small peptides present in urine and unlike other techniques such as liquid chromatography (LC), there is no requirement for continuous adaptation of conditions for optimal ionisation. This allows automated urinalysis which would be ideally suited to a diagnostic tool in the clinical setting.

Figure 6.1. ***Simple schematic of the CE-MS process*** *(adapted with kind permission from Mosaiques Diagnostics): 1. The CE electrode is inserted into the urine sample applying an electrical field and the sample is introduced into the capillary. 2. Analytes are separated according to ionic mobility 3. The capillary outlet is introduced to the electrospray ionisation chamber (ESI) 4. The ions produced are analysed by mass spectrometry.*



Sample preparation

For proteomic analysis 0.7 millilitres (mls) of urine was supplemented with 0.7 mls of 2M urea and 10mM NH40H containing 0.02% SDS. To remove proteins of higher molecular mass (for example albumin and immunoglobulin G) the sample was filtered using a Centrisart ultracentrifugation filter device (20 kDa (kilodalton) molecular weight cut-off; (Sartorius, Goettingen, Germany) at 2600xg until 1.1 mls of filtrate was obtained. The filtrate was then loaded onto a PD-10 desalting column (GE Healthcare, Munich, Germany) and equilibrated in 0.01% NH4OH in HPLC-grade water to decrease matrix effects by removing urea, electrolytes and salts and

also to enrich polypeptides. All samples were lyophilised and suspended in HPLC-grade water at 4^0C directly before CE-MS analysis.

CE-MS analysis

A P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, USA) was used coupled to a Micro-TOF (time-of-flight) MS (Bruker Daltonic, Bremen, Germany). A solution of 20% acetonitrile (Sigma-Aldrich, Taufkirchen, Germany) in HPLC-grade water supplemented with 0.94% formic acid (Sigma-Aldrich) was used as a running buffer. The ESI sprayer (Agilent Technologies, Palo Alto, USA) was grounded and the ion spray interface potential was set between -4.0 and -4.5 kV. Data acquisition and MS acquisition methods were automatically controlled by the CE via contact-close-relays. Spectra were accumulated every 3 seconds over a range of m/z 350 to 3000.

CE-MS data processing

Ion peaks were decluttered using MosaiquesVisu software which uses a probabilistic clustering algorithm and is a well-established method (371). Each polypeptide is characterised by its mass and migration time. Using reference data points the peptides were calibrated using local linear regression to isolate points of interest and discard ionic signal intensity that resulted from environmental factors (e.g. salt content). The resulting peak list from each sample analysis characterises each polypeptide by it's a) calibrated molecular mass b) calibrated CE migration time and c) normalised signal intensity. Raw peptide data was collated in a Microsoft database and disease-specific peptide models were generated using the support vector machine (SVM)-based MosCluster software. Classification of patient samples is achieved using the SVM-based multi-marker model. Here a peptide classifier is constructed, and each sample is given a score expressing the similarity of the patients' peptide profile to that of one identified to be specific to the disease. The classifier based test is constructed using an array of biomarkers. In general, it is not possible to reduce highly complex disease pathophysiology to a single marker (372). In addition, a single peptide marker is highly susceptible to the effects of biological variability.

Sequencing of peptides

Urine samples were analysed on a Dionex Ultimate 3000 RSLS nano flow system (Dionex, Camberly, UK). Elution was performed on an Acclaim PepMap C18 nano column with a linear gradient of 0.1% formic acid as solvent A versus 100% acetonitrile as solvent B starting at 5% to 50% over 100 minutes. The sample was ionised in positive ion mode using a Proxeon nano spray ESI source (Thermo Fisher Hemel, UK) and analysed in an Orbitrap Velos FTMS (Thermo Finnigan, Bremen, Germany). Data files were searched against the IPI human non-redundant database using the Open Mass Spectrometry Search Algorithm without any enzyme specificity. Calculated CE-migration time based on the number of basic amino acids was compared to the experimental migration time.

Statistical analysis

Estimates of sensitivity and specificity for the SVM-based disease specific peptide models are calculated by tabulating the number of correctly classified samples. MedCalc 8.1.1.0 (Mariakerke, Beligium) was used to calculate 95% CI's and ROC plot. The area under the ROC curve (AUC) was calculated as it provides a single measure of overall accuracy (373).

Incorporating previous plasma analysis: A previous, small feasibility study using CE-MS at Mosaiques-diagnostics established 392 plasma peptides that were differentially regulated between CRC (n=18) and normal (n=12) patients. From these peptides 143 amino acid sequences and 24 proteins were sequenced. Due to the low numbers of samples in this pilot study on urine there was not enough statistical power for adjusted p-value statistics. To compensate for this statistical limitation it was decided to only select those urinary peptide markers that could be corroborated with peptide markers found in plasma. Urinary peptides were selected with significant Wilcoxon rank sum p-values with a frequency distribution of 30% in at least one group.

Results

A total of 24 patient urinary samples were analysed: 12 with CRC, 6 with colonic adenomas and 6 normal controls (table 6.1). 5 of the CRCs were right sided and 8 were left sided. Dukes stage for the cancer cases was as follows: Dukes A – 1/12 (8.3%), Dukes B – 5/12 (41.7%), Dukes C – 6/12 (50%). Four adenoma cases were low risk and two were intermediate, all were found in the left colon.

Table 6.1. Study population characteristics

	Median age [years] (sd)	**Male (%)**	**Median BMI [kg/m$^{2]}$ (sd)**	**Smoker %**	**Alcohol intake [average u/wk]**
CRC	70 (9.4)	8/12 (66.7)	26.1 (2.0)	41.7	9.5
Adenoma	75 (4.9)	4/6 (66.7)	27 (3.9)	66.7	6
Control	67(7.4)	5/6 (83.3)	24.8 (2.9)	0	1.5

CRC peptide markers in urine

158 urinary peptide marker candidates with significant Wilcoxon p-values were detected, i.e. markers that were differently regulated in CRC compared with adenoma + control. Using tandem mass spectrometry 52 amino acids were isolated from the peptide markers (figure 6.2). The 19 proteins from which the amino acids are derived, are shown in table 6.2. The contour plots that contain the raw CE-MS data are shown in figure 6.3, here the crude differences in peptide composition between the groups can be visualised.

Figure 6.2. Identification chain from urinary peptides to proteins

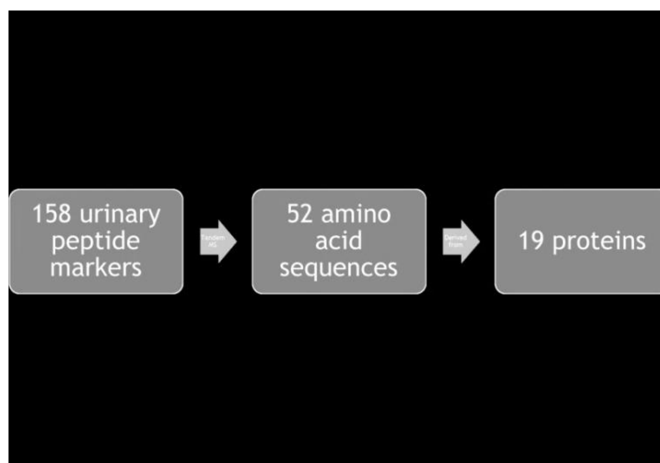
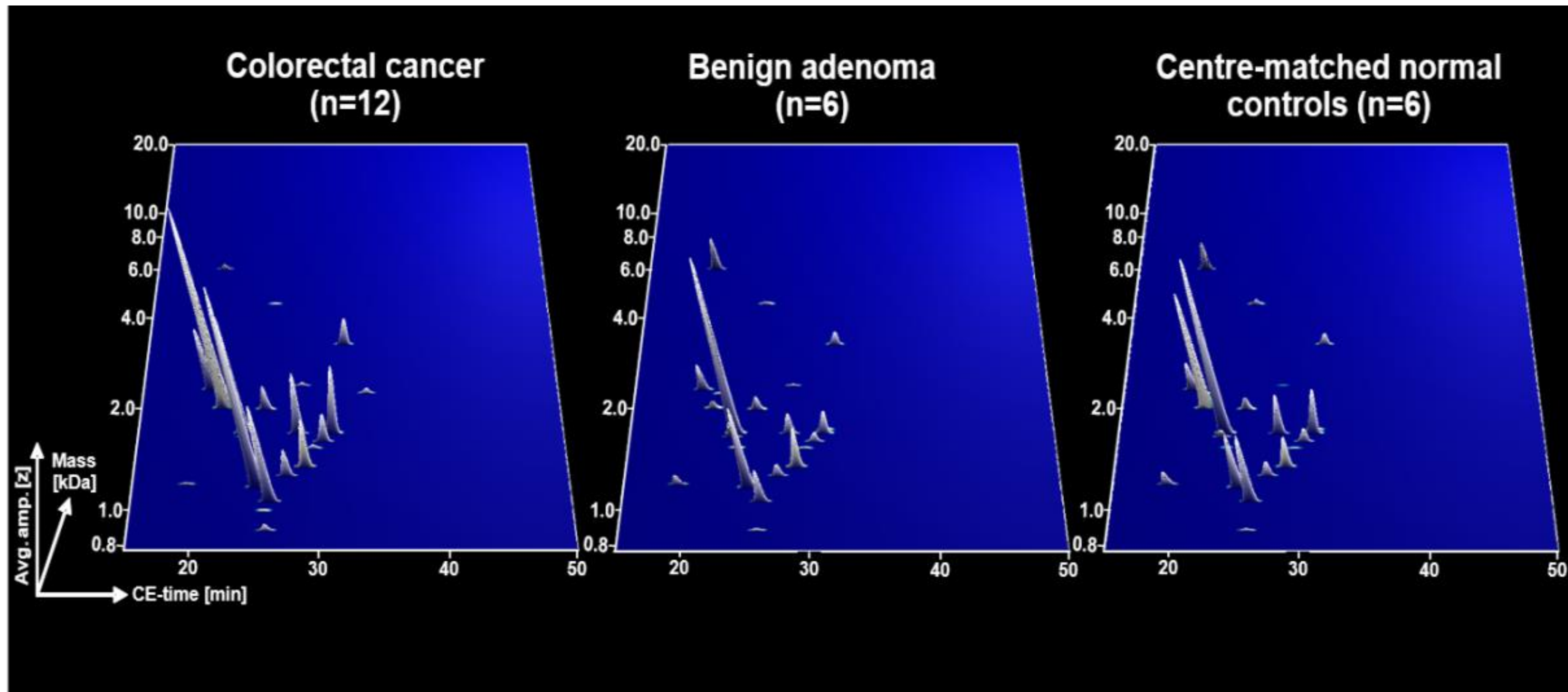


Table 6.2. Proteins from which the CRC peptide marker candidates in urine are derived *and number of peptide markers for each protein precursor.*

Protein name	**Number of peptide markers**
Collagen alpha-1(I) chain	22
Collagen alpha-1(III) chain	7
Polymeric immunoglobulin receptor	3
Uromodulin	3
Clusterin	2
Collagen alpha-1(II) chain	2
CD99 antigen	1
Collagen alpha-1(V) chain	1
Collagen alpha-1(XI) chain	1
Collagen alpha-1(XXVI) chain	1
Collagen alpha-2(I) chain	1
Collagen alpha-3(IX) chain	1
Fibrinogen alpha chain	1
Gelsolin	1
Heat shock protein beta-1	1
Protein S100-A9	1
Sodium/potassium-transporting ATPase subunit gamma	1
Xylosyltransferase 1	1
Zyxin	1

The most prominent protein was collagen alpha-1(I) chain with 22 corresponding peptide markers. Other key proteins are collagen alpha-1(III) with seven corresponding peptides, uromodulin (three peptides), polymeric immunoglobulin receptor (three peptides), clusterin (two peptides), collagen alpha-1(II) (two peptides). There were single peptide occurrences for zyxin, fibrinogen alpha chain, gelsolin, protein S100-A9, CD99 antigen, heat shock protein beta-1, sodium/potassium-transporting ATPase subunit gamma, xylosyltransferase 1 and five other collagen alpha chains.

Figure 6.3. Contour plots of the raw CE-MS data for CRC, adenoma and control *demonstrating the spectra of urinary polypeptides. CE migration time (x axis) is plotted against log molecular weight (y axis). Mean signal intensity is demonstrated by peak height.*



Performance of a CRC-specific peptide marker model

Both the complete peptide marker lists for plasma and for urine were screened to identify peptides with identical experimental mass, CE migration time and sequence identity. Ten peptides fulfilled these criteria. A second method to identify peptide markers common to both plasma and urine was also employed in order to compensate for the differences in protease composition present due to the different composition of each bodily fluid. Overlapping amino acid sequences were assessed between the two CRC groups. Frequent overlaps were seen in the collagen alpha-1(I) chain. This method, of identifying common peptide markers from the same linear region highlighted a total of 16 additional CRC urinary peptide marker candidates. In total 26 CRC urinary peptide markers were linked to peptide markers previously defined in CRC plasma samples. These 26 markers were used to generate a SVM-based, CRC specific multidimensional peptide marker panel. The complete list of peptide markers is shown in table 6.3.

Using the SVM-based classification model derived from the 26 surrogate CRC peptide markers, it was possible to differentiate the 12 CRC cases from 12 non-CRC cases with 100% sensitivity and 92% specificity, AUC 0.99 95% CI:0.84-1. These results are shown in table 6.4 and the corresponding ROC curve is shown in figure 6.4.

Peptide ID† (urine)	CE-MS characteristics		Distribution in normal controls		Distribution in CRC		Amino acid sequence and protein information		
	Exp. mass [Da]	CE-time [min]	Mean (SD)	Frequency	Mean (SD)	Frequency	Sequence‡	Protein symbol	Protein name
5675	911.435	25.8763	161 (191)	67	493 (517)	100	DGKTGPpGPA	COL1A1	Collagen alpha-1(I) chain
13992	1031.43	25.6586	0 (0)	0	31 (51)	42	NGDDGEAGKpG	COL1A1	Collagen alpha-1(I) chain
17694	1096.48	26.0757	4431 (4614)	100	6008 (3146)	92	ApGDRGEpGPp	COL1A1	Collagen alpha-1(I) chain
24502	1200.54	25.0266	3604 (4173)	83	10031 (3864)	100	KGDAGApGApGSQG	COL1A1	Collagen alpha-1(I) chain
25893	1223.57	19.3926	894 (1522)	67	138 (447)	17	DHEGTHSTKRG	FGA	Fibrinogen alpha chain
30575	1297.58	27.365	1018 (1438)	83	1820 (1080)	92	SpGSpGPDGKTGPp	COL1A1	Collagen alpha-1(I) chain
35339	1378.61	28.822	2018 (802)	100	3071 (923)	100	ApGDRGEpGPpGPAG	COL1A1	Collagen alpha-1(I) chain
46649	1563.7	29.4617	42 (94)	33	277 (173)	92	SpGSPGPDGKTGpPGPAG	COL1A1	Collagen alpha-1(I) chain
46883	1567.73	23.9151	9 (16)	33	0 (0)	0	VQEKQHPVPPPAQN	ZYX	Zyxin
51932	1636.74	30.251	890 (874)	67	1813 (1286)	75	GSpGSpGPDGKTGPpGPAG	COL1A1	Collagen alpha-1(I) chain
55315	1693.76	23.4788	472 (1019)	33	5 (17)	8	PpGPpGKNGDDGEAGKpG	COL1A1	Collagen alpha-1(I) chain
57265	1732.77	28.1753	2468 (1401)	83	3655 (1435)	100	WVGTGASEAEKTGAQEL	GSN	Gelsolin
57531	1737.78	31.0012	2788 (1814)	83	4057 (851)	100	DKGEpGGpGADGVPGKDGP	COL3A1	Collagen alpha-1(III) chain
57537	1737.78	23.7334	8584 (3818)	100	6390 (2046)	100	NDGAPGKNGERGGpGGpGp	COL3A1	Collagen alpha-1(III) chain
59645	1779.8	31.1464	321 (596)	50	0 (0)	0	GLTGSpGSpGPDGKTGPpGP	COL1A1	Collagen alpha-1(I) chain
72896	2055.94	25.438	641 (426)	100	1474 (673)	100	SGEpGApGSKGDTGAKGEpGPVG	COL1A1	Collagen alpha-1(I) chain
73697	2070.92	25.399	662 (386)	100	1134 (472)	100	GNSGEpGApGSKGDTGAKGEPGp	COL1A1	Collagen alpha-1(I) chain

73913	2076.95	21.7789	355 (738)	50	1252 (995)	92	GPpGPpGKNGDDGEAGKpGRpG	COL1A1	Collagen alpha-1(I) chain
74420	2085.93	22.0672	5831 (9863)	33	8401 (6652)	67	EGSpGRDGSpGAKGDRGETGPA	COL1A1	Collagen alpha-1(I) chain
86241	2303	22.2953	44 (80)	33	1 (5)	8	n.i.	---	---
87272	2319.07	33.817	0 (0)	0	317 (422)	42	AGppGEAGKPGEQGVpGDLGApGPSG	COL1A1	Collagen alpha-1(I) chain
90344	2377.1	20.7997	1644 (1275)	100	3474 (1622)	100	GKNGDDGEAGKPGRpGERGPpGpQ	COL1A1	Collagen alpha-1(I) chain
93417	2446.09	28.3726	5 (13)	17	185 (185)	75	ADGQpGAKGEpGDAGAKGDAGPpGPAGP	COL1A1	Collagen alpha-1(I) chain
130747	3359.58	31.8979	657 (716)	83	1509 (769)	100	pPGADGQPGAKGEpGDAGAKGDAGppGPAGPAGPPGPIG	COL1A1	Collagen alpha-1(I) chain
165081	4638.1	25.7822	284 (581)	50	25 (86)	8	n.i.	---	---
177971	6236.91	21.066	1375 (765)	83	242 (373)	42	n.i.	---	---

* Abbreviations: AA, amino acid; AUC, area under the curve; CE-MS, capillary electrophoresis mass spectrometry; Da, Dalton; n.i., not identified; min, minutes; SD, standard deviation.

† Peptide identification number.

‡ Lower case p and m indicates hydroxyproline and oxidized methionine.

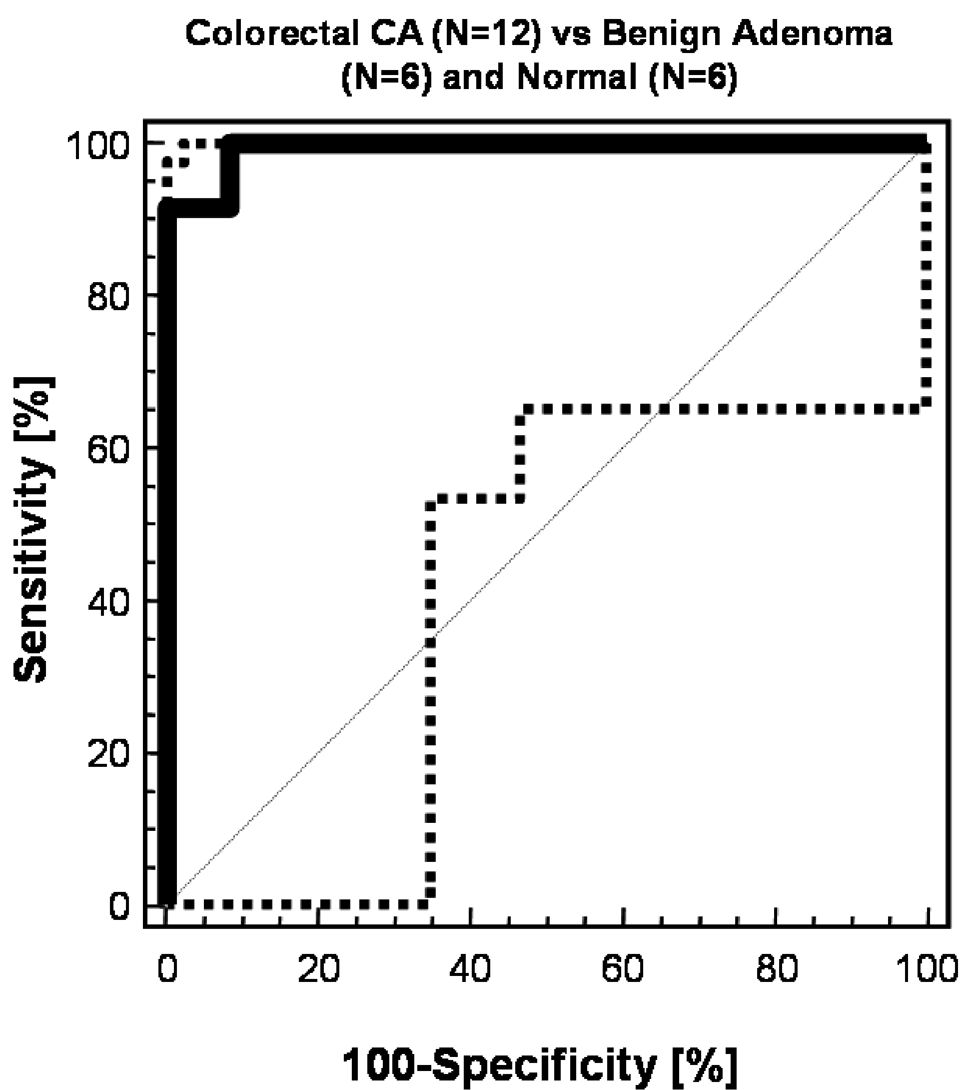
Table 6.3. CE-MS and sequence characteristics of the 26 CRC peptide marker candidates that were detected by CE-MS in plasma and urine *and that showed in both biofluids significant differences in their distribution by group wise comparison of CE-MS peptide profiles between CRC cases and non-cancer control*

Table 6.4 Classification of CRC from control using a 26-peptide urinary marker test

	Result
AUC	0.99* 95% CI (0.84-1)
Sensitivity	100
Specificity	0.92

**P=0.0001*

Figure 6.4. ROC curve for differentiating CRC from non-CRC using a 26 urinary peptide marker model *with corresponding AUC, sensitivity and specificity. The dotted lines either side of the true result line represent upper and lowert 95% CI's.*



6.5.2 Experiment two: Establishing the role of urinary-derived proteins in CRC.

Setting: Recruitment took place at UHCW and analysis was undertaken at the Clinical Sciences Research Laboratories within UHCW. Inclusion criteria: CRC cases recruited from experiment one, provided the tumour had been removed and was available for analysis. Exclusion criteria: Cases where no tissue specimen was available. Cases where patient had undergone chemotherapy, radiotherapy or any other cancer-related treatment prior to surgical excision of tumour.

Method

Patient characteristic data is shown in table 6.1 in the above section.

Antibody selection: The raw peptide data for both serum and urine analysis was interrogated to establish those with a direct sequence match in MEROPS (an online database for the classification of peptidases) and those with the highest number of cleavage sites in order to identify a single protein, or 'antigen' of interest. Nine showed potential, and are listed in table 6.5 together with corresponding antibody code:

Table 6.5 Nine antigens (proteases) which have a link with the raw urinary peptide data *based on highest number of cleavage sites.*

Target antigen	Antibody code
Cathepsin D	ab75852
Cathepsin E	ab36996
Cathepsin G	ab192793
Kalikrein 14	ab203226
Meprin alpha	ab107548
Meprin beta	ab204886
MMP3	ab52915
MMP7	ab205525
Plasminogen	ab174285

Secondly, the protease prediction data was reviewed using Proteasix, which is an open-source peptide-central tool which can predict the proteases involved in native peptide generation. Using this information, the protease selected for investigation was MEP1A. The rationale for this was three-fold: Firstly, it is connected to thymosin beta-4 which was identified in 4 peptide markers in plasma and has observed cleavage sites for MEP1A, as found on MEROPS. Secondly, thymosin beta-4 has a known role in CRC and finally, there has already been a link made between MEP1A and CRC in the literature.

Preparation of slides: The slides were formalin-fixed and paraffin-embedded at UHCW using a standardised protocol.

Optimising the antibody: The Human Protein Atlas (374) was interrogated, to investigate the current guidance on antibody dilutions and anticipated staining patterns for MEP1A. This was used as a guide for the serial dilution experiments that were conducted to optimise stain concentration. A MEP1A-specific primary antibody (at a dilution of 1:1000; Abcam, UK) was used for all IHC. This is a rabbit polyclonal to MEP1A, known to react with human tissue.

IHC: The method for IHC was carried out in accordance with the standard operating procedure documents produced by the UHCW tissue bank. However, consideration was made of specific manufacturer instructions regarding optimal dilution and antigen retrieval. The combination of primary antibody, dilution and detection system were validated on a series of test controls prior to use on the cases. Sections (taken from the same patients studied in experiment one), of colonic tumour (n=12) and adjacent normal tissue (n=12) were cut from stored blocks located in UHCW tissue bank. Each slice was reviewed by a Gastrointestinal Pathologist to ensure adequate tissue was collected prior to staining. Sections were cut at 3-4 μm and dried overnight in a 56^{0}C oven. A Leica polymer detection kit (Leica, UK) was used throughout.

- Pre-labelled slides were washed with distilled water.
- A hydrophobic pen was used to mark borders around the tissue sections, they were then placed in a metal moist chamber.

- Endogenous peroxidase activity was blocked using 100 µl/2 drops of Leica peroxidase block and left for 5 minutes.
- Two washes were carried out using Tris-buffered saline (TBS) for five minutes each. Sections were then incubated in Leica protein block 100µl for 20 mins, this reduces non-specific binding of primary and polymer.
- The protein block was drained off and a further 2 TBS washes were completed at five minutes each.
- Incubation in MEP1A at 1:800 dilution was then performed as determined by prior serial dilutions. Incubation with antibody was performed overnight in the fridge at 4^0C to help reduce the level of background staining.
- After overnight incubation, sections were washed in TBS for 10 minutes, twice. Leica post primary block was then applied to slides using 100µl and incubated for 30 minutes, this solution recognises the rabbit immunoglobulins, this step is followed by two washes with TBS for 5 minutes each.
- Sections were then incubated in Leica Novolink polymer solution using 100µl to ensure adequate coverage and incubated for 30 minutes followed by two washes with TBS for 5 minutes each time.
- Sections were incubated in Leica diaminobenzidine (DAB) working solution. This was constituted using 50µl of DAB chromagen in one ml of Novolink DAB substrate. Each slide was covered with two-three drops and left for five minutes, reaction with the peroxidase produces a visible brown precipitate at the antigen site.
- This step was followed by a further two washes with TBS for five minutes each time. Slides were then rinsed in distilled water and drained off.
- A counterstain of haematoxylin (brown colour) was applied for 30 seconds and slides were again washed in distilled water, followed by incubation in TBS for 2-3 minutes. Once left to dehydrate they were then ready for viewing.
- Results were interpreted using a light microscope and Panoramic viewer (3DHISTECH Ltd, Hungary), where digitally converted slides were displayed.

The 12 CRC and adjacent normal tissue slides were examined by a Gastrointestinal Pathologist at UHCW, who provided evaluation of the staining intensity using the Allred scoring system, the details of which are in table 6.6.

Table 6.6 Allred scoring system for stain intensity (375)

Score	Intensity
0	Negative (no staining of any nuclei at high magnification)
1	Weak (only visible at high magnification)
2	Moderate (readily visible at low magnification)
3	Strong (strikingly positive at low magnification)

Results

Study population characteristics are displayed in table 6.1 in the section above. There were 13 cancers from 12 patients (one synchronous cancer case). Tumour characteristics are listed in table 6.7 according to Dukes stage.

Table 6.7. Tumour site and stage

Case Number	Tumour Site	Dukes stage	TNM
1*	Rectum	A	T2N0M0
2	Rectum	A	T1N0M0
3	R. Colon	B	T3N0M0
4	R. Colon	B	T4N0M0
5	Rectum	B	T3N0M0
6	Rectum	B	T3N0M0
7	T. Colon	B	T3N0M0
8*	Caecum	C1	T3N2M0
9	R. Colon	C1	T3N1M0
10	Rectum	C1	T2N2M0
11	Sigmoid	C1	T4N1M0
12	Sigmoid	C1	T3N1M0
13	Rectum	C2	T4N2M0

*Synchronous tumours from same patient

Table 6.8 shows the staining score for each case and control. The stain intensity scores have been divided into four, corresponding to the tumour tissue, the epithelium of the adjacent normal tissue and the stroma in both CRC and normal tissue.

Table 6.8 IHC staining intensity score for MEP1A in CRC compared with normal tissue.

	Tumour stain intensity	**Epithelial stain intensity**	**Stromal stain intensity**	
Tissue type / Case number	**CRC**	**Normal**	**CRC**	**Normal**
1	3	2	1	0
2	1	2	1	0
3	3	3	3	0
4	2	2	0	0
5	3	2	1	0
6	3	3	3	1
7	3	2	0	0
8	3	3	0	0
9	3	2	1	0
10	No result	2	0	0
11	3	3	0	0
12	3	3	0	0
13	1	2	0	0

Antibody staining of the colonic tissue and adjacent normal tissue using IHC resulted in three main observations:

1. There were a greater proportion of cases with a stain intensity score of 3 (strong intensity) in the CRC group, compared with the normal group. In the CRC group 9/12 (75%) had a score of 3 and 5/13 (38%) in the normal group had a score of 3. This difference was not statistically significant ($z=1.069$ $p=>0.2$).
2. In 4 (33%) of the CRC cases MEP1A was concentrated in the advancing margin of the tumour, 2 of these cases were Dukes B and 2 Dukes C.
3. 6/13 (46.1%) CRC cases showed stromal staining with MEP1A compared with 1/13 (7.7%) of the normal cases. In most cases the stain intensity score was 1 in CRC.

Figure 6.5a-c shows digital images of cancer tissue and adjacent normal tissue with the difference in distribution of MEP1A illustrated. MEP1A staining is brown. The objective magnification is denoted in brackets. Figure 6.5a clearly shows strong epithelial staining in normal tissue compared with the surrounding stroma. In figure 6.5b strong stromal uptake of MEP1A in the cancer tissue is demonstrated. In the low power image in figure 6.5c, there is an accumulation of MEP1A stain seen at the advancing border of the cancer, this was observed in a third of cases.

Figure 6.5a. Digital image of IHC staining of MEP1A in the normal colonic tissue epithelium of case number 6. (x200). *There is intense MEP1A (brown) staining in the concentric epithelium (circled) compared with low intensity staining of the surrounding stroma (arrow)*

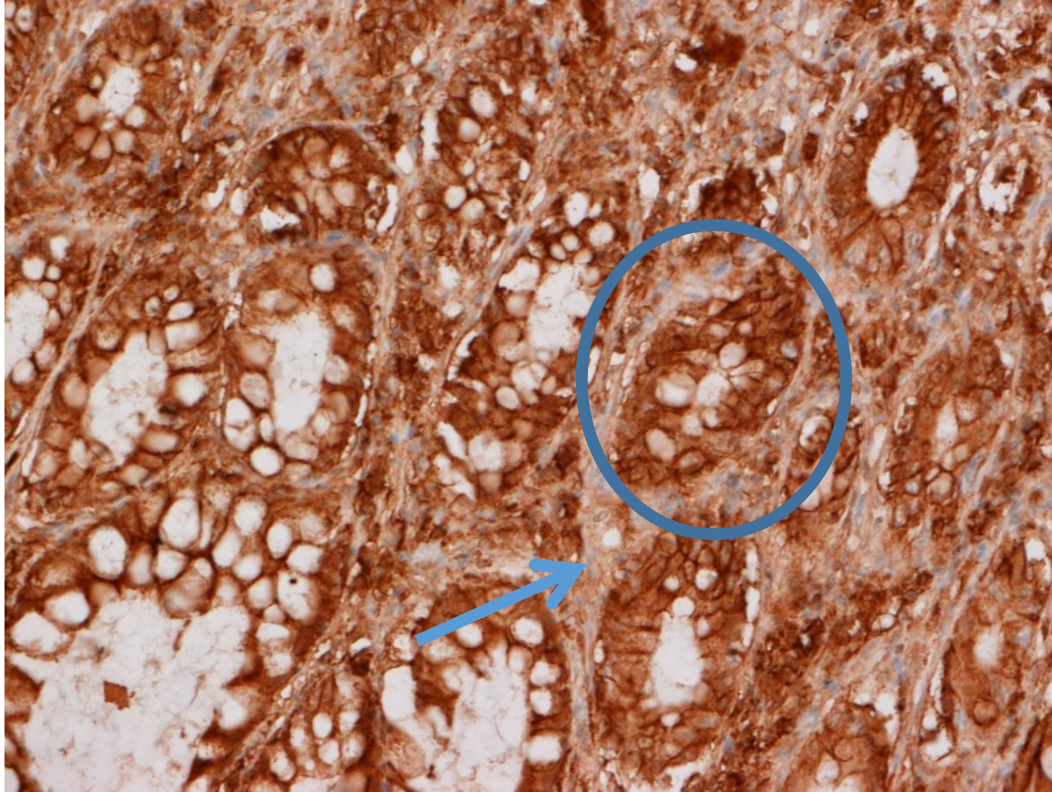


Figure 6.5b Digital image of IHC staining of MEP1A in the CRC tissue of case number 6 (x600). *The arrow highlights intense MEP1A accumulation (brown stain) in the stroma. This is in direct contrast to the reverse appearance seen in figure 6.5a.*

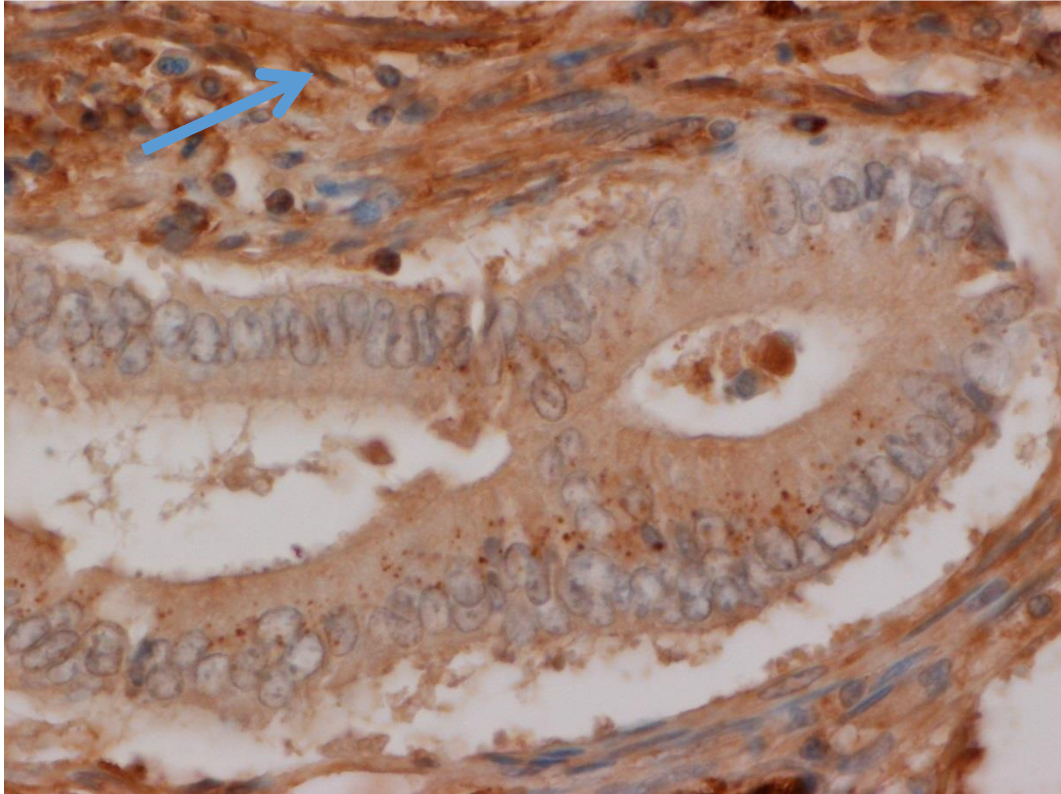
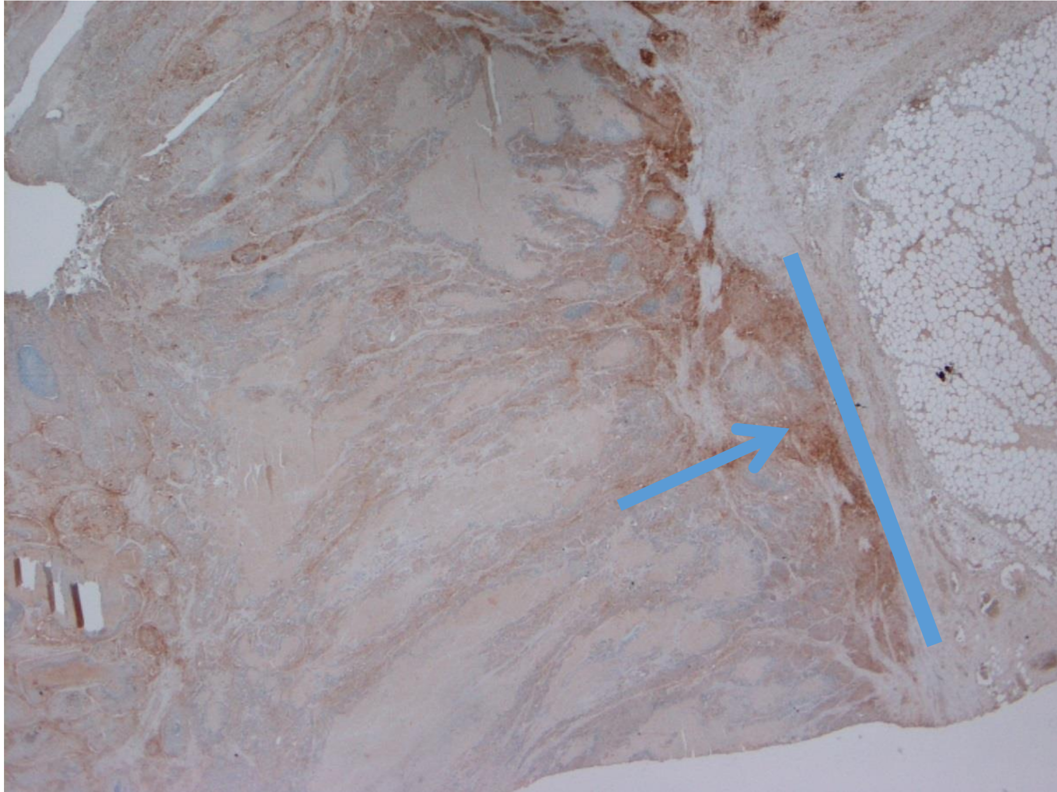


Figure 6.5c Digital image of IHC staining of MEP1A in the advancing margin of CRC tissue in case number 8. *(x10). Cancer tissue is seen to the left of the line, normal tissue is to the right. The deep brown staining at the junction of these tissues (arrow) represents MEP1A accumulation in the cancer tissue.*



6.6 Discussion

The feasibility of peptide based markers as a diagnostic tool in CRC and the subsequent identification of possible target proteins involved in tumourigenesis, were explored in this chapter. There have been a number of important findings that contribute to the body of evidence on proteomics in the detection of CRC.

Using CE-MS urinary peptide sequencing, 158 peptide markers were upregulated in the CRC samples, compared with control. A classification model based on 26 peptides common to both the plasma and urine of patients with CRC was proposed. In this proof-of-concept study, applying this model demonstrated high sensitivity (100%), specificity (92%) and overall test accuracy (AUC 0.99 95% CI 0.84-1) in correctly classifying CRC. This indicates the potential of peptide markers in the

diagnosis of CRC. In addition, this work is the first to utilise urine for CE-MS analysis, proving that it is feasible but also highly practical in the clinical setting given that CE-MS is user-friendly and portable. The suitability of urine as a diagnostic medium has been emphasised on several occasions. This work should stimulate further investigation into the capabilities of CE-MS in CRC detection using urine. Furthermore an added advantage of studying peptides, is the ability to exploit the gathered peptide data, in order to ascertain the proteins from which the peptides derive. Identification of these proteins may provide insight into their role in tumour formation and progression.

The sequenced peptides in this urinary study, were fragments of 19 proteins. The most prominent proteins were collagen alpha-1(I) and (III). Collagen alpha-1 (1) and (111) (COL1A1 and COL3A1) are members of the collagen family and are mainly expressed in extensible connective tissue such as skin and vessels. They have both been linked to CRC. A recent study by Zhang et al found increased COL1A1 in CRC tumour tissue and in metastatic lymph tissue and postulate that COL1A1 functions as an oncogene in CRC progression and is associated with metastatic disease (376). This may be due to the association of COL1A1 with the wingless type/planar cell polarity (WNT/PCP) pathway, although more work is needed to explore this further. Earlier work by Zou et al also demonstrated upregulation of COL1A1 in CRC using LC-MS/MS and its presence was associated with poorer survival (377). COL3A1 has also been linked to poorer prognosis in CRC, with Wang and colleagues showing upregulation of epithelial COL3A1, rather than stromal. Increase in tissue collagen was mirrored with an increase in serum COL3A1 also observed (378). To date, however there has not been replication of these findings in CRC, and both COL3A1 and COL1A1 have been linked to other cancers too, so their significance is still not clear (379-382).

Further utilisation of sequenced urinary peptide markers to better understand the pathology of disease, has been attempted by a small number of groups, but none in CRC. This is achieved by determining the presence or absence of target proteins in cancer tissue directly by using immunoassay techniques. One such technique that can be employed is immunohistochemistry (IHC).

IHC is a method of identifying the presence and distribution of antigen proteins in the cells of a tissue section, by using an antibody to bind specifically to the antigen in question. The immunogenic response to cancer has been well described and IHC has been utilised successfully in previous studies of ovarian, lung and prostate cancer (383).

The protein profile in a patient with cancer is, in part, altered by the immune response of the host to that cancer. This is demonstrated by the identification of autoantibodies specific for a number of intracellular and cell-surface antigens, that are detected in the serum of affected patients (250). Identifying antigenic biomarker has so far, been under-researched but is a promising area, as the immune system can amplify the immune-mediated response making antigenic tumour proteins more detectable.

In view of the advantages discussed, the peptide marker results from experiment one were used as a platform to identify their proteins of origin, and establish whether they can be linked to the cancer site directly. This may suggest a role in the pathology of CRC.

A single protein called MEP1A, was targeted at the cancer tissue site. Meprin, a metalloendopeptidase, is a member of the astacin family and was originally discovered in 1986 by Bond et al, as a protease highly expressed at the kidney brush border membrane (in mouse models) (384). It was subsequently found in the human intestinal membrane by Sterchi et al (385). There are two homologous subunits to meprin: alpha and beta.

Results from the second experiment in this chapter, found that MEP1A stain uptake was distributed differently between the CRC and normal tissue. Firstly, there were a greater proportion of CRC cases with stromal uptake compared with in normal tissue (46.1% and 7.7% respectively). In addition, there was a greater proportion of cases with a stain intensity score of strong in the tumour tissue (75%) compared with the normal epithelium (38%). This did not hold statistical significance. Another striking feature in a third of cases, was the aggregation of MEP1A at the advancing margin of the tumour.

MEP1A has been linked to a number of cancers including breast and also infections of the renal and urinary tract (384, 386). There has been a small body of work looking at the role of MEP1A in the pathogenesis of CRC. In 1999 Lottaz et al hypothesised that the basolateral secretion of MEP1A by colonic cancer exposed the extracellular matrix components and other stromal elements to increased proteolytic potential in the vicinity of tumour cells (387). Using IHC and also Western Blot they found a differentiated human colon carcinoma cell line endogenously expressed MEP1A. In 2011 a group headed by Lottaz again, examined stage 3 and 4 CRC and normal tissue and found MEP1A to be differently distributed in disease compared with control. In normal tissue they found MEP1A in epithelial cells, but in colonic cancer cells the protease was secreted in a non-polarised way which led to accumulation and activation in the tumour stroma (388). They postulate that extracellular proteolytic events may promote tumour progression by disrupting physical barriers. MEP1A staining in adenomatous tissue and in hepatic metastases was low.

Another group led by Rosmann in 2002 also found accumulation of MEP1A in the colonic tumour stroma and overall more MEP1A in tumour tissue compared with control. They postulated that activation of this protease occurred by plasmin and plasminogen (389). More recently, in 2016, Wang et al suggest that MEP1A plays a crucial role in CRC carcinogenesis and in particular it is involved in tumour migration and invasion. Their study showed elevated MEP1A expression in human CRC tissue, they also demonstrated in mice that when you knock-out the expression of MEP1A, cell proliferation and invasion was blocked, this supports the theory that MEP1A has a negative impact on tumorigenesis and progression (390).

The findings reported in this chapter, are supportive of the hypothesis that MEP1A has a role in the progression of CRC and its presence may signify a more aggressive tumour. This may hold importance in future treatment regimes for CRC. The presence or absence or MEP1A in the tumour margin may be useful as an indication of poorer outcome and thus aid in decisions regarding chemotherapeutic agents by providing a role in risk-stratification. This could be established in a follow up study linking those with MEP1A in the tumour margin to survival outcome.

Limitations

There were limitations to the work in this experiment. Firstly there are drawbacks to peptide marker selection using only adjusted Wilcoxon rank sum p-values. This was calculated due to the small sample size. However, this was compensated for by restricting the classifier model to only 26 peptide markers that also demonstrated association to CRC in plasma as well as urine. Further work involving a case-control study on a larger patient set is warranted to validate these tentative results on a larger scale, to ensure overfitting of data has not occurred. As long-term follow up data was not available on these patients, the significance of MEP1A on prognosis can only be hypothesised.

6.6.1 Returning to the research questions initially posed

1. Do patients with CRC have a different urinary peptidome to controls?

   The results from the first experiment have demonstrated a different urinary peptide profile in CRC compared with control. A test based on a combined plasma and urinary peptide marker model was trialled and demonstrated that CRC could be differentiated from control with a high degree of accuracy AUC (0.99), sensitivity (1.00) and specificity (0.92).
2. Which are the key proteins coded for by the peptide sequences detected in the urine of patients with CRC?

   The two most prominent proteins were COL1A1 and COL3A1.
3. Which corresponding proteins are present at the tumour site and what is their role in carcinogenesis?

   Nine proteins of interest were detected (table 6.5). MEP1A was found to have stronger stain intensity in the CRC tissue compared with normal tissue. Stromal stain uptake was present in CRC but absent in normal tissue. These findings support the evidence in the literature of a link between MEP1A and CRC that possibly indicates a more aggressive tumour phenotype.

6.7 Conclusion

This pilot study is the first to establish a urinary peptide marker profile for CRC which was differentiated from control with a high degree of accuracy and could form the basis of a diagnostic test.

Using an immunoassay, increased MEP1A in CRC tissue was demonstrated when compared with normal epithelium and suggests a link with more aggressive tumour pathology. This supports the current body of published literature on MEP1A. Based on the available research, there are likely to be other proteins with a similar pattern of distribution and derivation of a panel of proteins with high tumour and stromal uptake may give a more accurate prediction of tumour behaviour. This may have important implications for CRC treatment strategies and prognostication. Furthermore, the role and activity of a further eight proteases need to be investigated in CRC tissue.

There is huge scope for further work based on the results presented in this chapter, this exemplifies the expanse of protein-based work that is required in CRC to better understand its link with the proteome.

CHAPTER 7:

Discussion

7.1 Introduction

CRC is common and is a significant cause of morbidity and mortality across the world. The burden of this disease is rising in the UK. There is a lifetime risk of 1 in 19 for women, and 1 in 14 for men of developing CRC. Currently, overall five-year survival is approximately 60% (391). Five-year survival improves to nearly 90% for early stage disease, but unfortunately only the minority of patients present with early cancer (9). In comparison to many other cancers, funding into CRC research is high, with Cancer Research UK investing £35 million between 2016-2017 alone (392). Despite advances in the biology and genetics of CRC that have been achieved as a result of research funding, survival has not substantially improved in the last decade. In consideration of these factors, there is a need to re-evaluate how CRC is detected to establish areas for improvement and set achievable targets for future research.

The overall theme of this thesis was to explore non-invasive methods in the earlier detection of CRC. Firstly, the current TWW referral pathway for cases of suspected CRC was investigated, to evaluate whether this route to diagnosis was effective and improved outcome for patients. Secondly, the use of both urinary VOCs and peptides as biomarkers for the detection of CRC, was explored through experimental work. The aim was to understand their role in CRC detection and evaluate whether they represent a real option in the future for diagnosis of this disease.

7.2 The detection of CRC using the TWW referral pathway

The lower GI TWW referral pathway was devised in the year 2000 by the DOH, to streamline referrals for suspected CRC. There is increasing emphasis on utilising this pathway by NHS England. The National Cancer Strategy has set out ambitious aims in reducing cancer incidence and improving mortality rates (5).

Despite the drive for increased TWW usage by primary care, there are few measures in place for the re-evaluation of this referral pathway against key performance indicators. The Government initially strived for '*greater than 90% of cancer to be diagnosed*' using TWW pathways (144). The document reporting this statement has since been archived, and a target percentage is not quoted in the latest cancer strategy document.

The systematic review and meta-analysis in Chapter Three, was the first to be conducted with the aim of filling gaps in knowledge regarding diagnostic rates and outcomes for patients investigated via the TWW pathway. This systematic review reported a low yield of CRC of 7.7%. However, this was consistent with the cancer conversion rate for all cancers referred via any TWW pathway of 7.4%. Perhaps the most important finding of this review, was that no improvement in stage at diagnosis of CRC in the TWW cohort was found, when compared with stage at diagnosis via non-TWW outpatient routes. In the TWW group, 46.7% presented at Dukes C or above, in such cases survival drops substantially compared with earlier stage disease. This finding suggests that the TWW pathway is not having an effect on reducing mortality from CRC. In addition, the majority of patients had no significant lower GI pathology diagnosed, with 54.6% of patients found to have a normal colonic examination. An increased public awareness and lowering the thresholds for referral from primary care, have resulted in an increase in referral rates via TWW without an increase in the diagnosis of CRC (295). The latest figures from 2016-2017 showed an 8.7% rise in annual referrals overall for the TWW pathway (148).

The results of this systematic review indicate that the current triage of outpatients with lower GI symptoms is not effective in detecting those with CRC and has low sensitivity in detecting those with early stage disease. This raises important questions about whether the financial investment and clinical demands of the TWW pathway can be justified. Moreover, there is no rationale according to the results of this review, that the move to shorten diagnostic timeframes to 28 days is going to improve cancer stage or mortality (391).

In order to address the disappointing survival figures in the UK in comparison to Europe, there has to be a continued focus on improved pre-emptive medicine in CRC diagnosis, to increase the proportion of patients presenting at an early stage of disease. A symptom-based approach to diagnosis that forms the basis of the TWW referral criteria needs to be superseded by better risk stratification. There is currently on-going debate into how this is achieved.

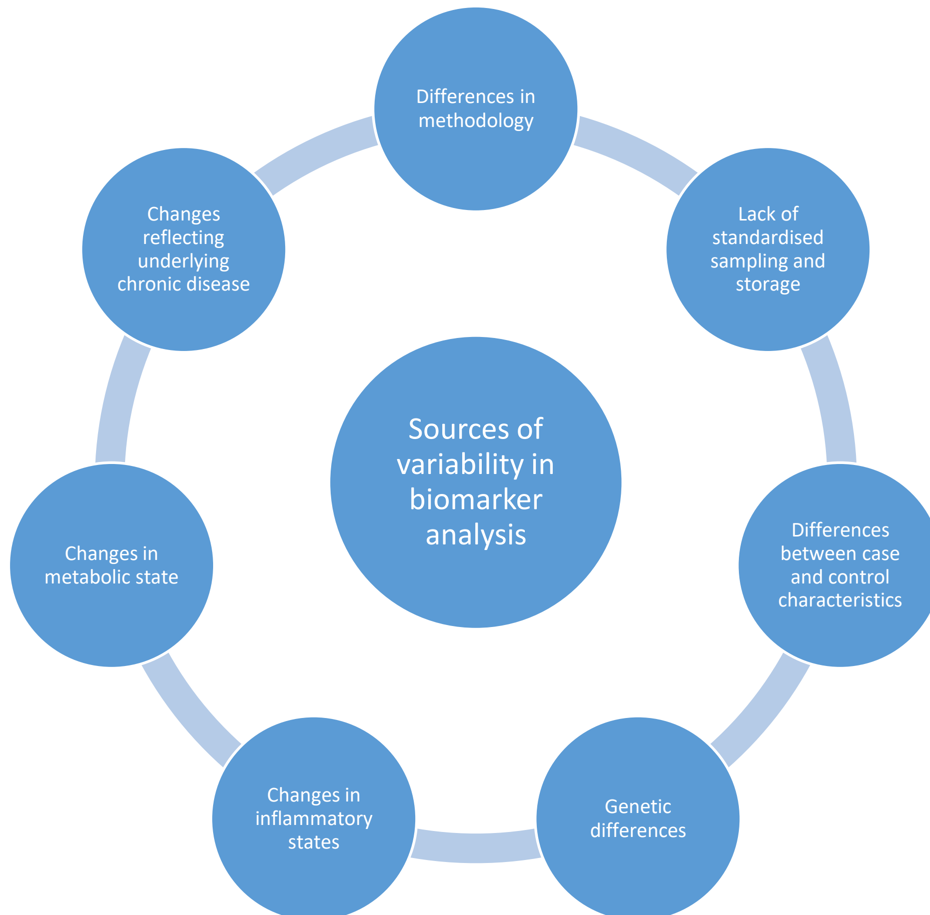
One possibility is the use of biomarkers, in combination with symptoms, to target more patients earlier in the course of the disease. For example, the most recently published NHS England guidance on TWW pathways in CRC has recommended the more widespread use of the faecal immunochemical test (FIT) for low-risk symptoms (391). This signifies a move to incorporate objective tests more widely, but there is a paucity of high quality options.

7.3 Biomarker development

Current biomarkers used in CRC are RAS and RAF mutation testing and mismatch repair gene deficiency. They serve to risk stratify the patient after diagnosis and provide predictions of response to certain treatments (393), but there is currently no detection-based biomarker in use.

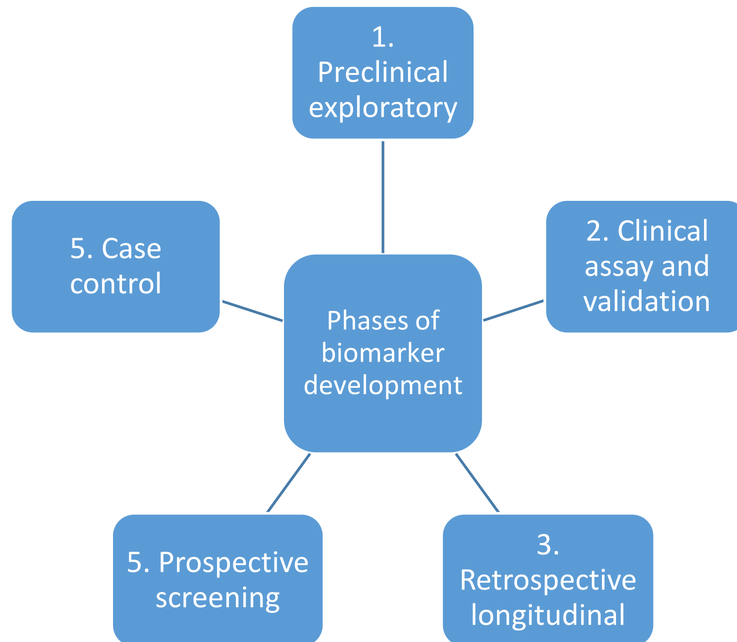
Devising a new biomarker is challenging. Major difficulties relate to the multitude of competing influences on these markers as a result of constant flux in the host and environment. The factors influencing biomarkers are shown in figure 7.1 and highlight the challenges of developing a disease-specific marker in an environment of constant change.

Figure 7.1 Schematic of the factors influencing potential biomarkers



Regardless of their target or intended use, biomarkers must also satisfy a number of conditions before they can be considered for clinical practice. Five phases of biomarker development have been proposed by Pepe and Barker and are listed in figure 7.2 (394). These five phases emphasise that in order for a biomarker to gain widespread use, it needs to be validated in prospective, well-controlled clinical studies including patients of varying age, sex and ethnicity and transcending multiple health-care sites, where regional variations in patient populations are accounted for. Furthermore the process must be tightly standardised at each step.

Figure 7.2 Five phases of biomarker development. *As described by Pepe and Barker (394).*



An important focus of this thesis was to evaluate some of the latest advances in biomarker technology in the detection of CRC, given the uncertainty surrounding the role of emerging biomarkers in the diagnostic pathway of CRC. Therefore, Chapters Four, Five and Six, explored the use of two novel biomarker groups.

Urine was the sole medium used for the experimental work due its many advantages. Patient satisfaction is high: over 90% of patients agreed to urine sampling, this compares to the uptake of the FOBT which is around 60%. It is readily available, shows stability in the medium term and reflects the internal environment of the body without the need for invasive sampling. It has been postulated that the urinary proteome is less complex than other bodily fluids such as blood (369). As a consequence, this may simplify analysis.

7.4 VOCs in the detection of CRC

The use of VOC-based disease detection has gained momentum over the last few decades and there has been a plethora of small studies examining the VOC profile of a variety of bodily fluids and linking this to disease detection. But there has been very little focussed research on the analysis and storage of samples, yet both play a crucial role in the validity of the results. Chapter Four provides the first detailed account of the salient considerations when analysing VOCs and will act as a framework for future researchers. The findings from the work examining the effect of storage on VOCs resulted in three key recommendations:

1. No changes to the environment or sampling methodology should be made during an experiment. Minor changes have been demonstrated to confound results.
2. Following sample collection, transfer to -80^{0}C storage should occur within a minimum of 12 hours.
3. Samples should be stored at -80^{0}C for a maximum of 12 months prior to analysis.

A comprehensive investigation of urinary VOC analysis in the detection of CRC was performed in Chapter Five. The results were unique in view of the size of the symptomatic population studied, but also because this was the first urinary VOC research involving the BCSP population. Analysis utilised machine learning models which are currently at the forefront of medical diagnostics research. The results indicated a high degree of accuracy in separation of CRC from non-CRC using FAIMS (AUC 0.98 95% CI: 0.93-1) in the FOBT positive population. Test accuracy dropped considerably when attempts were made to group CRC with other diagnoses and compare with different groups (AUC 0.56-0.64). This supports the hypothesis that CRC has a unique VOC signature. This is the first study to demonstrate the feasibility of urinary VOCs as a screening tool in the BCSP. The results confirm that there is potential in an alternative means of CRC detection in this group and provides a benchmark for future studies.

Adenomas represent a pre-malignant process, therefore identification and removal prevents cancer development in the majority of CRC cases. This is particularly

pertinent to the BCSP population in view of the fact that adenoma detection rates are far higher, with around 50% of patients found to have adenomas in the BCSP population compared with around 15% in the average population (340). Using FAIMS, CRC was correctly classified from low, intermediate and high risk adenomas with test accuracy ranging from 0.83-0.92 (95% CI: 0.66-1).

GC-IMS, a new analytical instrument in the detection of VOCs for medical diagnostics, was also used to investigate the BCSP cohort as a proof of concept study. GC-IMS was used to classify CRC cases from control. There was a high test accuracy (AUC 0.82 95% CI: 0.67-0.97), sensitivity was 0.80 (0.44-0.97) and specificity was also high at 0.83 (0.63-0.95) for correctly classifying cases based on VOC profile, this showed consistency with the FAIMS results.

This is one of only a handful of studies to demonstrate the application of this tool in the medical field for diagnostic purposes, and the first to look at CRC. The results were promising and should stimulate more interest in its application on further, larger-scale studies. Not only did this instrument demonstrate feasibility in detecting cancer from control, GC-IMS have several advantages as a clinical tool. It is simple to use, meaning specialist training is not required. Also, it is portable, thus in theory, could be transported to the clinical area of need for instant analysis. Finally, due to the IMS component, it has the technology to isolate and identify the chemical compounds within the urine sample. This has the potential to broaden current understanding of CRC pathogenesis and provide better encoding of features, resulting in more targeted VOC detection. The chemical library required to perform this step is currently in the development stage.

7.5 Proteomics in the detection of CRC

There is still currently very little understanding of the influence of proteomics on CRC or of the key proteins involved in carcinogenesis (392). In contrast to the high expectation, the impact of proteomics over the last two decades has been modest. This is mainly due to the unpredictable and complex nature of the proteome, which makes it exponentially more difficult to map than the genome (245).

The work in Chapter Six produced a number of interesting and promising results. The feasibility of urinary peptide markers as a test for CRC was investigated and was the first to demonstrate the utility of CE-MS in the peptide profiling of CRC. CE-MS is a highly sensitive modality and has several features that are suited to the clinical field including ease of use and fast analysis. Both these attributes are important in a busy clinical environment, as laboratory-based technology is often unable to meet the requirements of a medical setting. Using a 26-peptide test model, classification of CRC from control was high, with AUC 0.99 95% CI: 0.84-1, sensitivity was 100% and specificity was 92%., Based on these peptide markers, nineteen proteins were sequenced with COL1A1 AND COL1A3 showing prominence. In addition, using the raw peptide data to cross-reference with online protein databases, nine proteases were highlighted as having a potential link with CRC. These were MEP1A, MEP1B, MMP3, Cathepsin, G, Cathepsin E, Cathepsin D, Plasminogen, MMP7 and Kallikrein.

Using IHC, the role of MEP1A in the pathology of CRC was investigated. The results contributed to a small body of evidence on the topic (387-389). There were three main findings:

1. A greater proportion of CRC cases showed strong stain intensity of MEP1A compared with the adjacent normal tissue (75% versus 38%), suggesting MEP1A has a role in CRC.
2. In a third of CRC cases, MEP1A was seen to accumulate at the advancing margin of the tumour.
3. 46.1% of CRC cases demonstrated stromal staining with MEP1A compared with 7.7% of normal tissue.

Stromal migration of proteins and the consequent association with more advanced disease and resistance to treatment has been described in CRC (395, 396). But it is hard to make firm conclusions about this observation as studies have also demonstrated heterogeneity within the same tumour. For example, in one study patients could be simultaneously classified into multiple subgroups using a 30-gene list depending on the region analysed from the same tumour (397). The accumulation of MEP1A at the advancing edge of the tumour was an interesting

finding, not described before in the literature. One could postulate that this reinforces the theory of MEP1A proteolytic activity assisting in tumour cell migration into neighbouring structures.

Overall, the findings support the hypothesis that MEP1A has a role in the progression of CRC. Several features seen here may signify a more aggressive tumour type and could have implications in treatment regimes and risk – stratification. More evidence will need to be gained and will be aided by the improvements in analysis of formalin-fixed, paraffin-embedded tissues that have occurred in recent years.

The heterogeneity of the CRC proteome is considerable, with differences observed depending on the colonic tumour site and whether there are metastases or not (392). As mentioned, in research by previous groups, there has been heterogeneity seen within different sections from the same tumour specimen (392). It may be that the key functionality of proteome mapping in CRC is as an adjunct to other biomarkers and further research into a step-wise biomarker panel algorithm would be welcomed. In particular, a study combining investigation of other proteins that show increased stromal activity in cancer tissue may provide a more accurate prediction of tumour behaviour.

7.6 The future of biomarker research in CRC

Ultimately, there is a need to integrate with other research communities to share strategies and data more effectively. Most studies on biomarker research have been undertaken by individual investigators, focussed on single biomarker. This limits the strength and generalisability of the results. The current lack of approval and integration of biomarkers into clinical practice exemplifies the challenges that have already been discussed earlier in this chapter.

Multi-phase analysis is likely to be the most useful way of utilising biomarkers for maximum accuracy, i.e. using a set of biomarkers as part of a diagnostic algorithm that takes into account other factors too. This type of analysis will occur via more collaboration between groups who have access to different equipment and expertise.

There has already been a move to utilise FIT more broadly in the TWW population. It has been suggested that FIT could be incorporated into a more accurate assessment model by utilising other patient factors, such as age, sex, BMI and smoking status in predicting likelihood of disease (398). This is an interesting idea and challenges the population-based approach to most biomarker tests. This concept of risk stratification according to patient characteristics should be investigated by groups testing new biomarkers at the outset. It could provide a new level of understanding of individual patient risk and also prioritise investigations for those deemed at higher risk. Although the TWW pathway risk-stratifies patients according to symptoms, age is the only other factor taken into account. Yet, as described in Chapter Two other factors, such as diet, medication history and smoking status have an important effect on CRC risk.

Multi-institutional teamwork via participation in collaborative studies heightens both analytical and clinical validation, for example, proteomeXchange (399) is a consortium, which aims to coordinate mass spectrometry proteomics data. It also provide the most economical use of patient samples through standardisation of methodology which maximises validity and reduces bias.

Biomarker collaborative networks are in place, such as the National Cancer Institute Early Detection Research Network (400). To date, for CRC they have reported on 13 biomarkers, 12 of which are proteins. None currently have FDA approval, but results have the potential to create greater impact when collaboration occurs. Furthermore the Human Protein Atlas project has merged a partnership between the public and private sector to share data on tissue distribution and subcellular localisation of candidate biomarkers (374). Finally, the Global Alliance for Genomics and Health (401) is an international collaboration which provides information on genomics data sharing within the ethical, legal and technical boundaries. This provides reassurance to researchers who want to share knowledge, without compromising restrictions set by their local ethics board.

A forum for sharing experience and expertise on the analytical instruments used in both VOC and peptide analysis would be welcomed. This could include advice

when sampling errors occur, encourage the sharing of results and most importantly serve to form collaborations between groups.

Earlier this year, a highly instructive position paper by Lawler and colleagues was published. This outlines research gaps and provided recommendations to improve outcomes in CRC. This paper was written in collaboration with Bowel Cancer UK who have recently created a research initiative to guide upcoming researchers to concentrated funding in key areas according to need (392). This approach demonstrates a progressive strategy for change.

In regard to biomarkers, '*Metaphorically speaking, the 'water' is everywhere, but little is yet ready to drink'* (402).

7.7 Recommendations and further research

There is a lack of widely accepted prognostic and predictive biomarkers in CRC. Urine-based biomarkers could revolutionise CRC diagnostics with low cost and maximum accessibility.

The work in this thesis could be expanded further in several ways:

1. A larger-scale study of the detection of CRC from control. Using urinary VOCs from the symptomatic population would provide more robust data on the performance of a classifier test and would allow examination of the VOC profile by tumour stage, which was not possible in this thesis.
2. Use GC-IMS to identify the individual chemicals excreted as urinary VOCs in CRC and explore their role in carcinogenesis.
3. Explore the significance of the remaining eight proteins that have been implicated here in the pathology of CRC and use IHC to map their distribution in cancer tissue compared with control.

7.8 Limitations

There are several limitations to discuss. Firstly, the systematic review provided high quality evidence summarising the literature on the TWW referral pathway but the results, where high heterogeneity were found, must be deliberated with caution.

Attempts were made to ascertain the cause of heterogeneity but stratification by sample size did not result in any significant reduction.

There must be consideration of the small sample size in both the VOC and peptide biomarker studies. The machine learning algorithms that were used to analyse both sets of data always risk the possibility of overfitting of the data. This was minimised as much as possible by using a cross-validation technique. Research results must demonstrate repeatability and be validated on a larger scale before any strong conclusions can be drawn.

7.9 Final remarks

There needs to be a culture change from the fragmented approach to elucidating diagnostics in CRC. Large scale studies are required to be conducted where possible, in conjunction with research partners. This will enhance recruitment numbers, but also diversify the patient population to enable more generalisation of results. This is likely to happen through shared databases. Such work has already started to happen locally, driven by the FAMISHED research group, whereby co-ordination between four local trusts has already been set-up with preliminary results on VOC analysis within the next 12 months.

In conclusion, the work in this thesis draws together two key themes in the improvement of CRC detection. Firstly, purely symptom- based criteria or results of a single biomarker test are not the solution to improving CRC detection. Rather a biomarker profile that may be derived from multiple origins, e.g. gaseous, protein and genomic, should be incorporated into a risk-stratification algorithm of an individual patient, based on key characteristics. Secondly, the recommendation that collaboration is essential to validate pilot studies on a large scale and to share knowledge on methods and novel results.

Bibliography

1. Office for National Statistics: Mortality Statistics-Deaths registered in England and Wales Published online2014 [Available at: http://www.ons.gov.uk/ons/publications/all-releases.html?definition=tcm%3A77-27475].

2. (WHO). WHO. Leading cause of death in Europe: fact sheet Copenhagen: WHO Regional Office for Europe; 2012 [Available at: https://data.euro.who.int/hfadb].

3. Cancer Research UK: Bowel Cancer Diagnosis and Treatment Statistics 2015 [Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/diagnosis-and-treatment#heading-Zero].

4. Office for National Statistics: Cancer Registration Statistics, England:2014 2014 [https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014].

5. Department of Health: Improving Outcomes: A Strategy for Cancer 2011 [Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213785/dh_123394.pdf].

6. Department of Health:The likely impact of earlier diagnosis of cancer on costs and benefits to the NHS. London2011.

7. Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, et al. Cancer Survival in Australia,Canada, Denmark,Norway,Sweden, and the UK. Lancet. 2011;377(9760):127-138

8. Morris EJ, Sandin F, Lambert PC, Bray F, Klint A, Linklater K, et al. A population-based comparison of the survival of patients with colorectal cancer in England, Norway and Sweden between 1996 and 2004. Gut. 2011;60(8):1087-93.

9. Cancer Research UK 2014 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#ref-2].

10. Gatta G, Capocaccia R, Sant M, Bell CM, Coebergh JW, Damhuis RA, et al. Understanding variations in survival for colorectal cancer in Europe: a EUROCARE high resolution study. Gut. 2000;47(4):533-8.

11. Keane MG, Johnson GJ. Early diagnosis improves survival in colorectal cancer. Practitioner. 2012;256(1753):15-8, 2.

12. Bowel Cancer UK. Media: Latest news 2012 [Available from: http://www.bowelcanceruk.org.uk/media-centre/latest-news/cost-of-bowel-cancer-16-billion/].

13. Feldmen M FL, Brandt LJ. Gastrointestinal and Liver Disease. Pathophysiology/diagnosis/managment. Tenth ed: Elsevier Health Sciences; 2015.

14. Shepherd NA, Warren BF, Williams GT, Greenson JK, Lauwers GY, Novelli MR, et al. Morson and Dawson's gastrointestinal pathology. Available from: http://0-

onlinelibrary.wiley.com.pugwash.lib.warwick.ac.uk/book/10.1002/9781118399668
Connect to Wiley e-book.

15. Drake RL MJ. Anatomy of the Colon and Rectum 2016 [Available at: https://basicmedicalkey.com/anatomy-of-the-colon-and-rectum/].

16. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, Hemmelgarn BR. Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2009;7(12):1272-8.

17. Rex DK. Colonoscopy: a review of its yield for cancers and adenomas by indication. Am J Gastroenterol. 1995;90(3):353-65.

18. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer. 1975;36(6):2251-70.

19. Leslie A, Carey FA, Pratt NR, Steele RJ. The colorectal adenoma-carcinoma sequence. Br J Surg. 2002;89(7):845-60.

20. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.

21. Shinya H, Wolff WI. Morphology, anatomic distribution and cancer potential of colonic polyps. Ann Surg. 1979;190(6):679-83.

22. Hoff G, Sauar J, Vatn MH, Larsen S, Langmark F, Moen IE, et al. Polypectomy of adenomas in the prevention of colorectal cancer: 10 years' follow-up of the Telemark Polyp Study I. A prospective, controlled population study. Scand J Gastroenterol. 1996;31(10):1006-10.

23. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. N Engl J Med. 1992;326(10):658-62.

24. O'Brien MJ, Winawer SJ, Zauber AG, Gottlieb LS, Sternberg SS, Diaz B, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology. 1990;98(2):371-9.

25. Bos JL. ras oncogenes in human cancer: a review. Cancer Res. 1989;49(17):4682-9.

26. Riggins GJ, Kinzler KW, Vogelstein B, Thiagalingam S. Frequency of Smad gene mutations in human cancers. Cancer Res. 1997;57(13):2578-80.

27. Cairns SR, Scholefield JH, Steele RJ, Dunlop MG, Thomas HJ, Evans GD, et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). Gut. 2010;59(5):666-89.

28. Johnson CD, Chen MH, Toledano AY, Heiken JP, Dachman A, Kuo MD, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med. 2008;359(12):1207-17.

29. Anatomy, Histology, Embryology, and Developmental Anomalies of the Small and Large Intestine. 10th ed. United States of America: Elsevier Saunders; 2016.

30. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med. 2003;348(10):919-32.

31. Burt RW, DiSario JA, Cannon-Albright L. Genetics of colon cancer: impact of inheritance on colon cancer risk. Annu Rev Med. 1995;46:371-9.

32. Bisgaard ML, Fenger K, Bulow S, Niebuhr E, Mohr J. Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. Hum Mutat. 1994;3(2):121-5.

33. Nielsen M, Morreau H, Vasen HF, Hes FJ. MUTYH-associated polyposis (MAP). Crit Rev Oncol Hematol. 2011;79(1):1-16.

34. Vasen HF, Moslein G, Alonso A, Aretz S, Bernstein I, Bertario L, et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP). Gut. 2008;57(5):704-13.

35. Jagelman DG, DeCosse JJ, Bussey HJ. Upper gastrointestinal cancer in familial adenomatous polyposis. Lancet. 1988;1(8595):1149-51.

36. Half E, Bercovich D, Rozen P. Familial adenomatous polyposis. Orphanet J Rare Dis. 2009;4:22.

37. Gardner EJ. A genetic and clinical study of intestinal polyposis, a predisposing factor for carcinoma of the colon and rectum. Am J Hum Genet. 1951;3(2):167-76.

38. Snover D, Ahnen D, Burt R, et al. In: Bozman, FT., Carneiro F, Hruban R, et al., editors. Serrated polyps of the colon and rectum and serrated (“hyperplastic”) polyposis. . Pathology and genetics Tumours of the digestive system [Internet]. 2010.

39. Rex DK, Ahnen DJ, Baron JA, Batts KP, Burke CA, Burt RW, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol. 2012;107(9):1315-29; quiz 4, 30.

40. Moreira L, Balaguer F, Lindor N, de la Chapelle A, Hampel H, Aaltonen LA, et al. Identification of Lynch syndrome among patients with colorectal cancer. Jama. 2012;308(15):1555-65.

41. Bonadona V, Bonaiti B, Olschwang S, Grandjouan S, Huiart L, Longy M, et al. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. Jama. 2011;305(22):2304-10.

42. Jarvinen HJ, Aarnio M, Mustonen H, Aktan-Collan K, Aaltonen LA, Peltomaki P, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology. 2000;118(5):829-34.

43. Midgley R, Kerr D. Colorectal cancer. Lancet. 1999;353(9150):391-9.

44. Mahapatra AK eSoccO-lsathwsaipsch. Staging of colorectal cancers. 2008. Available at: www.sgpgi.ac.in/path/seminars/ccastage.html

45. Segnan N PJ, von Karsa L, editors. European guidelines for quality assurance in colorectal cancer screening and diagnosis. Luxembourg; 2010.

46. Cancer Research UK. Bowel Cancer incidence statistics [www.cancerresearchuk.org/health-professional/cancer-statistics-by-type/bowel-cancer/incidence].

47. Karahalios A, English DR, Simpson JA. Weight change and risk of colorectal cancer: a systematic review and meta-analysis. Am J Epidemiol. 2015;181(11):832-45.

48. Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L, et al. Carcinogenicity of consumption of red and processed meat. Lancet Oncol. 2015;16(16):1599-600.

49. Park Y, Hunter DJ, Spiegelman D, Bergkvist L, Berrino F, van den Brandt PA, et al. Dietary fiber intake and risk of colorectal cancer: a pooled analysis of prospective cohort studies. Jama. 2005;294(22):2849-57.

50. Aune D, Chan DS, Lau R, Vieira R, Greenwood DC, Kampman E, et al. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. Bmj. 2011;343:d6617.

51. Kim YI, Mason JB. Nutrition chemoprevention of gastrointestinal cancers: a critical review. Nutr Rev. 1996;54(9):259-79.

52. Terry P, Giovannucci E, Michels KB, Bergkvist L, Hansen H, Holmberg L, et al. Fruit, vegetables, dietary fiber, and risk of colorectal cancer. J Natl Cancer Inst. 2001;93(7):525-33.

53. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. Am J Epidemiol. 1998;148(1):4-16.

54. Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. Jama. 2008;300(23):2765-78.

55. Fedirko V, Tramacere I, Bagnardi V, Rota M, Scotti L, Islami F, et al. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. Ann Oncol. 2011;22(9):1958-72.

56. van der Schee MP. The volatile metabolome and microbiome in pulmonary and gastro-intestinal disease. Holland: Amsterdam; 2015.

57. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. PLoS One. 2010;5(1):e8578.

58. Major G, Spiller R. Irritable bowel syndrome, inflammatory bowel disease and the microbiome. Curr Opin Endocrinol Diabetes Obes. 2014;21(1):15-21.

59. Giongo A, Gano KA, Crabb DB, Mukherjee N, Novelo LL, Casella G, et al. Toward defining the autoimmune microbiome for type 1 diabetes. Isme j. 2011;5(1):82-91.

60. Haines A, Metz G, Dilawari J, Blendis L, Wiggins H. Breath-methane in patients with cancer of the large bowel. Lancet. 1977;2(8036):481-3.

61. Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM, Fan TJ, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science. 2012;338(6103):120-3.

62. de Vos WM, de Vos EA. Role of the intestinal microbiome in health and disease: from correlation to causation. Nutr Rev. 2012;70 Suppl 1:S45-56.

63. Winters MD, Schlinke TL, Joyce WA, Glore SR, Huycke MM. Prospective case-cohort study of intestinal colonization with enterococci that produce extracellular superoxide and the risk for colorectal adenomas or cancer. Am J Gastroenterol. 1998;93(12):2491-500.

64. Huycke MM, Abrams V, Moore DR. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. Carcinogenesis. 2002;23(3):529-36.

65. Hardy RG, Meltzer SJ, Jankowski JA. ABC of colorectal cancer. Molecular basis for risk factors. Bmj. 2000;321(7265):886-9.

66. Lippman SM, Benner SE, Hong WK. Cancer chemoprevention. J Clin Oncol. 1994;12(4):851-73.

67. Rothwell PM, Wilson M, Elwin CE, Norrving B, Algra A, Warlow CP, et al. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. Lancet. 2010;376(9754):1741-50.

68. Arber N, Eagle CJ, Spicak J, Racz I, Dite P, Hajer J, et al. Celecoxib for the prevention of colorectal adenomatous polyps. N Engl J Med. 2006;355(9):885-95.

69. Velayos FS, Loftus EV, Jr., Jess T, Harmsen WS, Bida J, Zinsmeister AR, et al. Predictive and protective factors associated with colorectal cancer in ulcerative colitis: A case-control study. Gastroenterology. 2006;130(7):1941-9.

70. Rubin DT, Lashner BA. Will a 5-ASA a day keep the cancer (and dysplasia) away? Am J Gastroenterol. 100. United States2005. p. 1354-6.

71. Shaukat A, Scouras N, Schunemann HJ. Role of supplemental calcium in the recurrence of colorectal adenomas: a metaanalysis of randomized controlled trials. Am J Gastroenterol. 2005;100(2):390-4.

72. Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Practice Parameters Committee of the American College of Gastroenterology. Am J Gastroenterol. 2000;95(11):3053-63.

73. Chung M, Lee J, Terasawa T, Lau J, Trikalinos TA. Vitamin D with or without calcium supplementation for prevention of cancer and fractures: an updated meta-analysis for the U.S. Preventive Services Task Force. Ann Intern Med. 2011;155(12):827-38.

74. Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, et al. Folic acid for the prevention of colorectal adenomas: a randomized clinical trial. Jama. 2007;297(21):2351-9.

75. Meyskens FL, Jr., McLaren CE, Pelot D, Fujikawa-Brooks S, Carpenter PM, Hawk E, et al. Difluoromethylornithine plus sulindac for the prevention of sporadic colorectal adenomas: a randomized placebo-controlled, double-blind trial. Cancer Prev Res (Phila). 2008;1(1):32-8.

76. Pedersen TR, Berg K, Cook TJ, Faergeman O, Haghfelt T, Kjekshus J, et al. Safety and tolerability of cholesterol lowering with simvastatin during 5 years in the Scandinavian Simvastatin Survival Study. Arch Intern Med. 1996;156(18):2085-92.

77. Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford JD, Cole TG, et al. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. N Engl J Med. 1996;335(14):1001-9.

78. Nelson H, Petrelli N, Carlin A, Couture J, Fleshman J, Guillem J, et al. Guidelines 2000 for colon and rectal cancer surgery. J Natl Cancer Inst. 2001;93(8):583-96.

79. Bonjer HJ, Hop WC, Nelson H, Sargent DJ, Lacy AM, Castells A, et al. Laparoscopically assisted vs open colectomy for colon cancer: a meta-analysis. Arch Surg. 2007;142(3):298-303.

80. Jackson TD, Kaplan GG, Arena G, Page JH, Rogers SO, Jr. Laparoscopic versus open resection for colorectal cancer: a metaanalysis of oncologic outcomes. J Am Coll Surg. 2007;204(3):439-46.

81. Excellence NIfHaC. Colorectal cancer: diagnosis and management. 2011 [Available at: www.nice.org.uk/guidance/cg131].

82. Guerrera F, Mossetti C, Ceccarelli M, Bruna MC, Bora G, Olivetti S, et al. Surgery of colorectal cancer lung metastases: analysis of survival, recurrence and re-surgery. J Thorac Dis. 2016;8(7):1764-71.

83. Simmonds PC. Palliative chemotherapy for advanced colorectal cancer: systematic review and meta-analysis. Colorectal Cancer Collaborative Group. Bmj. 2000;321(7260):531-5.

84. Roth AD, Delorenzi M, Tejpar S, Yan P, Klingbiel D, Fiocca R, et al. Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. J Natl Cancer Inst. 2012;104(21):1635-46.

85. Lutz ST, Chow EL, Hartsell WF, Konski AA. A review of hypofractionated palliative radiotherapy. Cancer. 2007;109(8):1462-70.

86. Excellence NIfHaC. Cetuximab and panitumumab for previously untreated metastatic colorectal cancer. (TA439). 2017 [Available from: www.nice.org.uk/guidance/ta439].

87. Cancer Research UK. Cancer drugs 2017 [Available at: www.cancerresearchuk.org/about-cancer/cancer-in-general/treatment/cancer-drugs].

88. Moreno CC, Mittal PK, Sullivan PS, Rutherford R, Staley CA, Cardona K, et al. Colorectal Cancer Initial Diagnosis: Screening Colonoscopy, Diagnostic Colonoscopy, or Emergent Surgery, and Tumor Stage and Size at Initial Presentation. Clin Colorectal Cancer. 2016;15(1):67-73.

89. Halligan S, Wooldrage K, Dadswell E, Kralj-Hans I, von Wagner C, Edwards R, et al. Computed tomographic colonography versus barium enema for diagnosis of colorectal cancer or large polyps in symptomatic patients (SIGGAR): a multicentre randomised trial. Lancet. 2013;381(9873):1185-93.

90. Atkin W, Dadswell E, Wooldrage K, Kralj-Hans I, von Wagner C, Edwards R, et al. Computed tomographic colonography versus colonoscopy for investigation of patients with symptoms suggestive of colorectal cancer (SIGGAR): a multicentre randomised trial. Lancet. 2013;381(9873):1194-202.

91. von Wagner C, Ghanouni A, Halligan S, Smith S, Dadswell E, Lilford RJ, et al. Patient acceptability and psychologic consequences of CT colonography compared with those of colonoscopy: results from a multicenter randomized controlled trial of symptomatic patients. Radiology. 2012;263(3):723-31.

92. Langan RC, Mullinax JE, Raiji MT, Upham T, Summers T, Stojadinovic A, et al. Colorectal cancer biomarkers and the potential role of cancer stem cells. J Cancer. 2013;4(3):241-50.

93. Macdonald JS. Carcinoembryonic antigen screening: pros and cons. Semin Oncol. 1999;26(5):556-60.

94. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. J Clin Oncol. 2006;24(33):5313-27.

95. Labianca R, Nordlinger B, Beretta GD, Mosconi S, Mandala M, Cervantes A, et al. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2013;24 Suppl 6:vi64-72.

96. Saito G, Sadahiro S, Kamata H, Miyakita H, Okada K, Tanaka A, et al. Monitoring of Serum Carcinoembryonic Antigen Levels after Curative Resection of Colon Cancer: Cutoff Values Determined according to Preoperative Levels Enhance the Diagnostic Accuracy for Recurrence. Oncology. 2017;92(5):276-82.

97. Lech G, Slotwinski R, Slodkowski M, Krasnodebski IW. Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. World J Gastroenterol. 2016;22(5):1745-55.

98. National Institute of Health Care Excellence. Faecal calprotectin diagnostic tests for inflammatory diseases of the bowel. (DG11) 2013 [Available at: www.nice.org.uk/guidance/DG11].

99. von Arnim U, Wex T, Ganzert C, Schulz C, Malfertheiner P. Fecal calprotectin: a marker for clinical differentiation of microscopic colitis and irritable bowel syndrome. Clin Exp Gastroenterol. 2016;9:97-103.

100. von Roon AC, Karamountzos L, Purkayastha S, Reese GE, Darzi AW, Teare JP, et al. Diagnostic precision of fecal calprotectin for inflammatory bowel disease and colorectal malignancy. Am J Gastroenterol. 2007;102(4):803-13.

101. Simon K. Colorectal cancer development and advances in screening. Clin Interv Aging. 2016;11:967-76.

102. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. Gut. 2014;63(2):317-25.

103. Vega P, Valentin F, Cubiella J. Colorectal cancer diagnosis: Pitfalls and opportunities. World J Gastrointest Oncol. 2015;7(12):422-33.

104. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol. 2005;23(3):609-18.

105. Samowitz WS, Curtin K, Ma KN, Schaffer D, Coleman LW, Leppert M, et al. Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. Cancer Epidemiol Biomarkers Prev. 2001;10(9):917-23.

106. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. N Engl J Med. 2003;349(3):247-57.

107. Sinicrope FA, Sargent DJ. Clinical implications of microsatellite instability in sporadic colon cancers. Curr Opin Oncol. 2009;21(4):369-73.

108. Elsaleh H, Iacopetta B. Microsatellite instability is a predictive marker for survival benefit from adjuvant chemotherapy in a population-based series of stage III colorectal carcinoma. Clin Colorectal Cancer. 2001;1(2):104-9.

109. Loughrey M QP, Shepard N. Standards and datasets for reporting cancers

Dataset for histopathological reporting of colorectal cancer. London; 2017. Contract No.: G049.

110. Popat S, Houlston RS. A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis. Eur J Cancer. 2005;41(14):2060-70.

111. Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, et al. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. N Engl J Med. 2001;344(16):1196-206.

112. Russo A, Bazan V, Iacopetta B, Kerr D, Soussi T, Gebbia N. The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. J Clin Oncol. 2005;23(30):7518-28.

113. Shulman K, Cohen I, Barnett-Griness O, Kuten A, Gruber SB, Lejbkowicz F, et al. Clinical implications of UGT1A1*28 genotype testing in colorectal cancer patients. Cancer. 2011;117(14):3156-62.

114. Rouits E, Charasson V, Petain A, Boisdron-Celle M, Delord JP, Fonck M, et al. Pharmacokinetic and pharmacogenetic determinants of the activity and toxicity of irinotecan in metastatic colorectal cancer patients. Br J Cancer. 2008;99(8):1239-45.

115. Wilson J JG. Principles and practice of screening for disease. Geneva; 1968.

116. Macafee DA, Scholefield JH. Antagonist: population based endoscopic screening for colorectal cancer. Gut. 2003;52(3):323-6.

117. A. M. Moving from guaiac faecal occult blood test (gFOBT) to a faecal immunochemical test for haemoglobin (FIT) in the bowel screening programme: A consultation. London; 2015.

118. Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med. 1993;328(19):1365-71.

119. Kewenter J, Brevinge H, Engaras B, Haglind E, Ahren C. Results of screening, rescreening, and follow-up in a prospective randomized study for detection of colorectal cancer by fecal occult blood testing. Results for 68,308 subjects. Scand J Gastroenterol. 1994;29(5):468-73.

120. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348(9040):1467-71.

121. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348(9040):1472-7.

122. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemoccult. Bmj. 1998;317(7158):559-65.

123. 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;103(6):1541-9.

124. The UK CRC Screening Pilot Evaluation Team. Evaluation of the UK Colorectal Cancer Screening Pilot. Edinburgh; 2003.

125. National Screening Committee. The UK NSC recommendation on Bowel Cancer screening in adults 2016 [Available at: http://legacy.screening.nhs.uk/bowelcancer].

126. Logan RF, 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;61(10):1439-46.

127. Walsh JM, Terdiman JP. Colorectal cancer screening: clinical applications. Jama. 2003;289(10):1297-302.

128. Labster. Gas Chromatography – mass spectrometry (GC-MS). [Available at: https://theory.labster.com/gcms]

129. Soares-Weiser K BJ, Duffy S, St John J, Smith A, Westwood M, Kleijnen J. Diagnostic accuracy and cost-effectiveness of faecal occult blood test (FOBT) used in screening for colorectal cancer: A systematic revew. York: The University of York; 2007.

130. 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.

131. Chambers JA, Callander AS, Grangeret R, O'Carroll RE. Attitudes towards the Faecal Occult Blood Test (FOBT) versus the Faecal Immunochemical Test (FIT) for colorectal cancer screening: perceived ease of completion and disgust. BMC Cancer. 2015;16(1):96.

132. Liles EG, Perrin N, Rosales AG, Feldstein AC, Smith DH, Mosen DM, et al. Change to FIT increased CRC screening rates: evaluation of a US screening outreach program. Am J Manag Care. 2012;18(10):588-95.

133. Steele RJ, McDonald PJ, Digby J, Brownlee L, Strachan JA, Libby G, et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity. United European Gastroenterol J. 2013;1(3):198-205.

134. van Rossum LG, van Rijn AF, Laheij RJ, van Oijen MG, Fockens P, van Krieken HH, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology. 2008;135(1):82-90.

135. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JCIY, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut. 2010;59(01):62-8.

136. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet. 2010;375(9726):1624-33.

137. Segnan N, Armaroli P, Bonelli L, Risio M, Sciallero S, Zappa M, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. J Natl Cancer Inst. 2011;103(17):1310-22.

138. Hoff G, Grotmol T, Skovlund E, Bretthauer M. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. Bmj. 2009;338:b1846.

139. Holme O, Loberg M, Kalager M, Bretthauer M, Hernan MA, Aas E, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. Jama. 2014;312(6):606-15.

140. Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. N Engl J Med. 2012;366(25):2345-57.

141. Atkin WS, Cuzick J, Northover JM, Whynes DK. Prevention of colorectal cancer by once-only sigmoidoscopy. Lancet. 1993;341(8847):736-40.

142. Kaminski MF, Bretthauer M, Zauber AG, Kuipers EJ, Adami HO, van Ballegooijen M, et al. The NordICC Study: rationale and design of a randomized trial on colonoscopy screening for colorectal cancer. Endoscopy. 2012;44(7):695-702.

143. Quintero E, Castells A, Bujanda L, Cubiella J, Salas D, Lanas A, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Engl J Med. 2012;366(8):697-706.

144. Department of Health: The NHS Cancer Plan 2000 [Archived:[Available at: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4014513.pdf].

145. Cancer Research UK:The Route to Diagnosis for Bowel Cancer 2016 [Available at: www.cancerresearchuk.org/the-route-to-diagnosis-for-bowel-cancer/

146. Hamilton W. Cancer diagnosis in primary care. Br J Gen Pract. 2010;60(571):121-8.

147. National Institute for Health and Care Excellence (NICE):Suspected Cancer (update) NG12. London: NICE; 2015.

148. Samuels M DD, McDonnell P. NHS England: Waiting Times for Suspected and Diagnosed Cancer Patients. 2016-17 annual report. 2017 [Available at: www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2017/06/Cancer-Waiting-Times-Annual-Report-201617-1.pdf].

149. Redaniel MT, Ridd M, Martin RM, Coxon F, Jeffreys M, Wade J. Rapid diagnostic pathways for suspected colorectal cancer: views of primary and secondary care clinicians on challenges and their potential solutions. BMJ Open. 2015;5(10):e008577.

150. Wang S, Ang HM, Tade MO. Volatile organic compounds in indoor environment and photocatalytic oxidation: state of the art. Environ Int. 2007;33(5):694-705.

151. Kansal A. Sources and reactivity of NMHCs and VOCs in the atmosphere: a review. J Hazard Mater. 2009;166(1):17-26.

152. Probert CS, Reade S, Ahmed I. Fecal volatile organic compounds: a novel, cheaper method of diagnosing inflammatory bowel disease? Expert Rev Clin Immunol. 2014;10(9):1129-31.

153. Kataoka H, Saito K, Kato H, Masuda K. Noninvasive analysis of volatile biomarkers in human emanations for health and early disease diagnosis. Bioanalysis. 2013;5(11):1443-59.

154. Arasaradnam RP, Westenbrink E, McFarlane MJ, Harbord R, Chambers S, O'Connell N, et al. Differentiating coeliac disease from irritable bowel syndrome by urinary volatile organic compound analysis--a pilot study. PLoS One. 2014;9(10):e107312.

155. Shirasu M, Touhara K. The scent of disease: volatile organic compounds of the human body related to disease and disorder. J Biochem. 2011;150(3):257-66.

156. Arasaradnam RP, Covington JA, Harmston C, Nwokolo CU. Review article: next generation diagnostic modalities in gastroenterology--gas phase volatile compound biomarker detection. Aliment Pharmacol Ther. 2014;39(8):780-9.

157. Lippi G, Cervellin G. Canine olfactory detection of cancer versus laboratory testing: myth or opportunity? Clin Chem Lab Med. 2012;50(3):435-9.

158. Sonoda H, Kohnoe S, Yamazato T, Satoh Y, Morizono G, Shikata K, et al. Colorectal cancer screening with odour material by canine scent detection. Gut. 2011;60(6):814-9.

159. Quignon P, Kirkness E, Cadieu E, Touleimat N, Guyon R, Renier C, et al. Comparison of the canine and human olfactory receptor gene repertoires. Genome Biol. 2003;4(12):R80.

160. Willis CM, Church SM, Guest CM, Cook WA, McCarthy N, Bransbury AJ, et al. Olfactory detection of human bladder cancer by dogs: proof of principle study. Bmj. 2004;329(7468):712.

161. Pizzini A, Filipiak W, Wille J, Ager C, Wiesenhofer H, Kubinec R, et al. Analysis of volatile organic compounds in the breath of patients with stable or acute exacerbation of chronic obstructive pulmonary disease. J Breath Res. 2018;12(3):036002.

162. Rocco G, Pennazza G, Santonico M, Longo F, Rocco R, Crucitti P, et al. Breathprinting and Early Diagnosis of Lung Cancer. J Thorac Oncol. 2018;13(7):883-94.

163. Princivalle A, Monasta L, Butturini G, Bassi C, Perbellini L. Pancreatic ductal adenocarcinoma can be detected by analysis of volatile organic compounds (VOCs) in alveolar air. BMC Cancer. 2018;18(1):529.

164. Silva CL, Passos M, Camara JS. Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers--a powerful strategy for breast cancer diagnosis. Talanta. 2012;89:360-8.

165. Wang C, Ke C, Wang X, Chi C, Guo L, Luo S, et al. Noninvasive detection of colorectal cancer by analysis of exhaled breath. Anal Bioanal Chem. 2014;406(19):4757-63.

166. de Meij TG, Larbi IB, van der Schee MP, Lentferink YE, Paff T, Terhaar Sive Droste JS, et al. Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: proof of principle study. Int J Cancer. 2014;134(5):1132-8.

167. van der Schee MP, Paff T, Brinkman P, van Aalderen WM, Haarman EG, Sterk PJ. Breathomics in lung disease. Chest. 2015;147(1):224-31.

168. Smith D, Spanel P. Pitfalls in the analysis of volatile breath biomarkers: suggested solutions and SIFT-MS quantification of single metabolites. J Breath Res. 2015;9(2):022001.

169. Kolakowski BM, D'Agostino PA, Chenier C, Mester Z. Analysis of chemical warfare agents in food products by atmospheric pressure ionization-high field asymmetric waveform ion mobility spectrometry-mass spectrometry. Anal Chem. 2007;79(21):8257-65.

170. Covington JA, van der Schee MP, Edge AS, Boyle B, Savage RS, Arasaradnam RP. The application of FAIMS gas analysis in medical diagnostics. Analyst. 2015;140(20):6775-81.

171. Arasaradnam RP, Ouaret N, Thomas MG, Quraishi N, Heatherington E, Nwokolo CU, et al. A novel tool for noninvasive diagnosis and tracking of patients with inflammatory bowel disease. Inflamm Bowel Dis. 2013;19(5):999-1003.

172. Persaud K, Dodd G. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. Nature. 1982;299(5881):352-5.

173. Karasek FW CR. Basic gas chromatography-mass spectrometry: Principles and techniques. Science E, editor. New York1988.

174. Allers M, Langejuergen J, Gaida A, Holz O, Schuchardt S, Hohlfeld JM, et al. Measurement of exhaled volatile organic compounds from patients with chronic obstructive pulmonary disease (COPD) using closed gas loop GC-IMS and GC-APCI-MS. J Breath Res. 2016;10(2):026004.

175. Mochalski P, Wiesenhofer H, Allers M, Zimmermann S, Guntner AT, Pineau NJ, et al. Monitoring of selected skin- and breath-borne volatile organic compounds emitted from the human body using gas chromatography ion mobility spectrometry (GC-IMS). J Chromatogr B Analyt Technol Biomed Life Sci. 2018;1076:29-34.

176. Lewis JM, Savage RS, Beeching NJ, Beadsworth MBJ, Feasey N, Covington JA. Identifying volatile metabolite signatures for the diagnosis of bacterial respiratory tract infection using electronic nose technology: A pilot study. PLoS One. 2017;12(12):e0188879.

177. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci U S A. 1971;68(10):2374-6.

178. Phillips M. Method for the collection and assay of volatile organic compounds in breath. Anal Biochem. 1997;247(2):272-8.

179. Jareno-Esteban JJ, Munoz-Lucas MA, Gomez-Martin O, Utrilla-Trigo S, Gutierrez-Ortega C, Aguilar-Ros A, et al. Study of 5 Volatile Organic Compounds in Exhaled Breath in Chronic Obstructive Pulmonary Disease. Arch Bronconeumol. 2017;53(5):251-6.

180. Besa V, Teschler H, Kurth I, Khan AM, Zarogoulidis P, Baumbach JI, et al. Exhaled volatile organic compounds discriminate patients with chronic obstructive pulmonary disease from healthy subjects. Int J Chron Obstruct Pulmon Dis. 2015;10:399-406.

181. Christiansen A, Davidsen JR, Titlestad I, Vestbo J, Baumbach J. A systematic review of breath analysis and detection of volatile organic compounds in COPD. J Breath Res. 2016;10(3):034002.

182. Dragonieri S, Schot R, Mertens BJ, Le Cessie S, Gauw SA, Spanevello A, et al. An electronic nose in the discrimination of patients with asthma and controls. J Allergy Clin Immunol. 2007;120(4):856-62.

183. Fens N, Zwinderman AH, van der Schee MP, de Nijs SB, Dijkers E, Roldaan AC, et al. Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma. Am J Respir Crit Care Med. 2009;180(11):1076-82.

184. Pavlou A, Turner AP, Magan N. Recognition of anaerobic bacterial isolates in vitro using electronic nose technology. Lett Appl Microbiol. 2002;35(5):366-9.

185. van der Schee MP FN, Buijze H et al. Diagnostic value of exhaled breath analysis in tuberculosis. Am J Respir Crit Care Med. 2012;185:a6510.

186. Dragonieri S, Pennazza G, Carratu P, Resta O. Electronic Nose Technology in Respiratory Diseases. Lung. 2017;195(2):157-65.

187. Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, O'Neill HJ. Volatile organic compounds in exhaled air from patients with lung cancer. Clin Chem. 1985;31(8):1278-82.

188. Phillips M, Gleeson K, Hughes JM, Greenberg J, Cataneo RN, Baker L, et al. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet. 1999;353(9168):1930-3.

189. Schallschmidt K, Becker R, Jung C, Bremser W, Walles T, Neudecker J, et al. Comparison of volatile organic compounds from lung cancer patients and healthy controls-challenges and limitations of an observational study. J Breath Res. 2016;10(4):046007.

190. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer. 2010;103(4):542-51.

191. Phillips M, Cataneo RN, Cruz-Ramos JA, Huston J, Ornelas O, Pappas N, et al. Prediction of breast cancer risk with volatile biomarkers in breath. Breast Cancer Res Treat. 2018.

192. Phillips M, Cataneo RN, Saunders C, Hope P, Schmitt P, Wai J. Volatile biomarkers in the breath of women with breast cancer. J Breath Res. 2010;4(2):026003.

193. Abaffy T, Duncan R, Riemer DD, Tietje O, Elgart G, Milikowski C, et al. Differential volatile signatures from skin, naevi and melanoma: a novel approach to detect a pathological process. PLoS One. 2010;5(11):e13813.

194. de Gennaro G, Dragonieri S, Longobardi F, Musti M, Stallone G, Trizio L, et al. Chemical characterization of exhaled breath to differentiate between patients with malignant plueral mesothelioma from subjects with similar professional asbestos exposure. Anal Bioanal Chem. 2010;398(7-8):3043-50.

195. van Hooren MR, Leunis N, Brandsma DS, Dingemans AC, Kremer B, Kross KW. Differentiating head and neck carcinoma from lung carcinoma with an electronic nose: a proof of concept study. Eur Arch Otorhinolaryngol. 2016;273(11):3897-903.

196. Sun X, Shao K, Wang T. Detection of volatile organic compounds (VOCs) from exhaled breath as noninvasive methods for cancer diagnosis. Anal Bioanal Chem. 2016;408(11):2759-80.

197. Qin T, Liu H, Song Q, Song G, Wang HZ, Pan YY, et al. The screening of volatile markers for hepatocellular carcinoma. Cancer Epidemiol Biomarkers Prev. 2010;19(9):2247-53.

198. De Vincentis A, Pennazza G, Santonico M, Vespasiani-Gentilucci U, Galati G, Gallo P, et al. Breath-print analysis by e-nose for classifying and monitoring chronic liver disease: a proof-of-concept study. Sci Rep. 2016;6:25337.

199. Turner C. Techniques and issues in breath and clinical sample headspace analysis for disease diagnosis. Bioanalysis. 2016;8(7):677-90.

200. Khalid T, Aggio R, White P, De Lacy Costello B, Persad R, Al-Kateb H, et al. Urinary Volatile Organic Compounds for the Detection of Prostate Cancer. PLoS One. 2015;10(11):e0143283.

201. Bernabei M, Pennazza G, Santonico M, Corsi C, Roscioni C, Paolesse R, et al.

202. Wang D, Wang C, Pi X, Guo L, Wang Y, Li M, et al. Urinary volatile organic compounds as potential biomarkers for renal cell carcinoma. Biomed Rep. 2016;5(1):68-72.

203. Capelli L, Taverna G, Bellini A, Eusebio L, Buffi N, Lazzeri M, et al. Application and Uses of Electronic Noses for Clinical Diagnosis on Urine Samples: A Review. Sensors (Basel). 2016;16(10).

204. Roine A, Saviauk T, Kumpulainen P, Karjalainen M, Tuokko A, Aittoniemi J, et al. Rapid and accurate detection of urinary pathogens by mobile IMS-based electronic nose: a proof-of-principle study. PLoS One. 2014;9(12):e114279.

205. Banday KM, Pasikanti KK, Chan EC, Singla R, Rao KV, Chauhan VS, et al. Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects. Anal Chem. 2011;83(14):5526-34.

206. Arasaradnam RP, McFarlane M, Daulton E, Skinner J, O'Connell N, Wurie S, et al. Non-invasive exhaled volatile organic biomarker analysis to detect inflammatory bowel disease (IBD). Dig Liver Dis. 2016;48(2):148-53.

207. Arasaradnam RP, McFarlane M, Daulton E, Westenbrink E, O'Connell N, Wurie S, et al. Non-invasive distinction of non-alcoholic fatty liver disease using urinary volatile organic compound analysis: early results. J Gastrointestin Liver Dis. 2015;24(2):197-201.

208. Arasaradnam RP, Wicaksono A, O'Brien H, Kocher HM, Covington JA, Crnogorac-Jurcevic T. Noninvasive Diagnosis of Pancreatic Cancer Through Detection of Volatile Organic Compounds in Urine. Gastroenterology. 2018;154(3):485-7.e1.

209. Casalinuovo IA, Pierro D, Bruno E, Francesco P, Coletta M. Experimental use of a new surface acoustic wave sensor for the rapid identification of bacteria and yeasts. Lett Appl Microbiol. 2006;42(1):24-9.

210. Gibson TD, Prosser O, Hulbert JN, Marshall R, Corcoran P, Lowery P, et al. Detection and simultaneous identification of micro-organisms from

headspace samples using an electronic nose. Sens Actuat B. 1997;4:413-22.

211. Chandiok S, Crawley BA, Oppenheim BA, Chadwick PR, Higgins S, Persaud KC. Screening for bacterial vaginosis: a novel application of artificial nose technology. J Clin Pathol. 1997;50(9):790-1.

212. Wilson AD, Lester DG. Use of an electronic-nose device for profiling headspace volatile

metabolites to rapidly identify phytopathogenic microbes. Phytopathology. 1997(87):S116.

213. Parry AD, Chadwick PR, Simon D, Oppenheim BA, McCollum CN. Detection ofβ-haemolytic streptococcal infections by analysis of leg ulcer odour. J. Wound Care1994. p. 134-7.

214. Milanowski M, Pomastowski P, Ligor T, Buszewski B. Saliva - Volatile Biomarkersand Profiles. Crit Rev Anal Chem. 2017;47(3):251-66.

215. Navaneethan U, Parsi MA, Lourdusamy V, Bhatt A, Gutierrez NG, Grove D, et al. Volatile organic compounds in bile for early diagnosis of cholangiocarcinoma in patients with primary sclerosing cholangitis: a pilot study. Gastrointest Endosc. 2015;81(4):943-9.e1.

216. Sivertsen SM, Bjorneklett A, Gullestad HP, Nygaard K. Breath methane and colorectal cancer. Scand J Gastroenterol. 1992;27(1):25-8.

217. Kokoszka J, Nelson RL, Swedler WI, Skosey J, Abcarian H. Determination of inflammatory bowel disease activity by breath pentane analysis. Dis Colon Rectum. 1993;36(6):597-601.

218. Pelli MA, Trovarelli G, Capodicasa E, De Medio GE, Bassotti G. Breath alkanes determination in ulcerative colitis and Crohn's disease. Dis Colon Rectum. 1999;42(1):71-6.

219. Shepherd SF, McGuire ND, de Lacy Costello BP, Ewen RJ, Jayasena DH, Vaughan K, et al. The use of a gas chromatograph coupled to a metal oxide sensor for rapid assessment of stool samples from irritable bowel syndrome and inflammatory bowel disease patients. J Breath Res. 2014;8(2):026001.

220. Walton C, Fowler DP, Turner C, Jia W, Whitehead RN, Griffiths L, et al. Analysis of volatile organic compounds of bacterial origin in chronic gastrointestinal diseases. Inflamm Bowel Dis. 2013;19(10):2069-78.

221. Ahmed I, Greenwood R, Costello Bde L, Ratcliffe NM, Probert CS. An investigation of fecal volatile organic metabolites in irritable bowel syndrome. PLoS One. 2013;8(3):e58204.

222. Garner CE, Smith S, de Lacy Costello B, White P, Spencer R, Probert CS, et al. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. Faseb j. 2007;21(8):1675-88.

223. Hockstein NG, Thaler ER, Lin Y, Lee DD, Hanson CW. Correlation of pneumonia score with electronic nose signature: A prospective study. Ann Otol Rhinol Laryngol. 2005;114(7):504-8.

224. Moore JG, Jessop LD, Osborne DN. Gas-chromatographic and mass-spectrometric analysis of the odor of human feces. Gastroenterology. 1987;93(6):1321-9.

225. Suarez F, Furne J, Springfield J, Levitt M. Insights into human colonic physiology obtained from the study of flatus composition. Am J Physiol. 1997;272(5 Pt 1):G1028-33.

226. Suarez FL, Springfield J, Levitt MD. Identification of gases responsible for the odour of human flatus and evaluation of a device purported to reduce this odour. Gut. 1998;43(1):100-4.

227. Rossi M, Aggio R, Staudacher HM, Lomer MC, Lindsay JO, Irving P, et al. Volatile Organic Compounds in Feces Associate With Response to Dietary Intervention in Patients With Irritable Bowel Syndrome. Clin Gastroenterol Hepatol. 2018;16(3):385-91.e1.

228. Covington JA, Wedlake L, Andreyev J, Ouaret N, Thomas MG, Nwokolo CU, et al. The detection of patients at risk of gastrointestinal toxicity during pelvic radiotherapy by electronic nose and FAIMS: a pilot study. Sensors (Basel). 2012;12(10):13002-18.

229. Bhatt A, Parsi MA, Stevens T, Gabbard S, Kumaravel A, Jang S, et al. Volatile organic compounds in plasma for the diagnosis of esophageal adenocarcinoma: a pilot study. Gastrointest Endosc. 2016;84(4):597-603.

230. Goldberg EM, Blendis LM, Sandler S. A gas chromatographic--mass spectrometric study of profiles of volatile metabolites in hepatic encephalopathy. J Chromatogr. 1981;226(2):291-9.

231. Xue R, Dong L, Zhang S, Deng C, Liu T, Wang J, et al. Investigation of volatile biomarkers in liver cancer blood using solid-phase microextraction and gas chromatography/mass spectrometry. Rapid Commun Mass Spectrom. 2008;22(8):1181-6.

232. Altomare DF, Di Lena M, Porcelli F, Trizio L, Travaglio E, Tutino M, et al. Exhaled volatile organic compounds identify patients with colorectal cancer. Br J Surg. 2013;100(1):144-50.

233. Haick H, Broza YY, Mochalski P, Ruzsanyi V, Amann A. Assessment, origin, and implementation of breath volatile cancer markers. Chem Soc Rev. 2014;43(5):1423-49.

234. van de Kant KD, van der Sande LJ, Jobsis Q, van Schayck OC, Dompeling E. Clinical use of exhaled volatile organic compounds in pulmonary diseases: a systematic review. Respir Res. 2012;13:117.

235. Batty CA, Cauchi M, Lourenco 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.

236. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, Westenbrink E, Hodges P, Thomas MG, et al. Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. PLoS One. 2014;9(9):e108750.

237. Wang C, Li P, Lian A, Sun B, Wang X, Guo L, et al. Blood volatile compounds as biomarkers for colorectal cancer. Cancer Biol Ther. 2014;15(2):200-6.

238. Westenbrink E, Arasaradnam RP, O'Connell N, Bailey C, Nwokolo C, Bardhan KD, et al. Development and application of a new electronic nose instrument for the detection of colorectal cancer. Biosens Bioelectron. 2015;67:733-8.

239. Amal H, Leja M, Funka K, Lasina I, Skapars R, Sivins A, et al. Breath testing as potential colorectal cancer screening tool. Int J Cancer. 2016;138(1):229-36.

240. Di Lena M, Porcelli F, Altomare DF. Volatile Organic Compounds (VOC) as new biomarkers for colorectal cancer: a review. Colorectal Dis. 2016.

241. Sinha R, Ahn J, Sampson JN, Shi J, Yu G, Xiong X, et al. Fecal Microbiota, Fecal Metabolome, and Colorectal Cancer Interrelations. PLoS One. 2016;11(3):e0152126.

242. Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP. The state of molecular biomarkers for the early detection of lung cancer. Cancer Prev Res (Phila). 2012;5(8):992-1006.

243. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Primers. 2015;1:15065.

244. The Human Proteome Organization. The Human Proteome Project 2016 [Available at: https://hupo.org/human-proteome-project].

245. Herrmann PC, Liotta LA, Petricoin EF, 3rd. Cancer proteomics: the state of the art. Dis Markers. 2001;17(2):49-57.

246. Mullner S, Neumann T, Lottspeich F. Proteomics--a new way for drug target discovery. Arzneimittelforschung. 1998;48(1):93-5.

247. Wittke S, Schiffer E, Bauer HW. [Capillary electrophoresis coupled to mass spectrometry for proteome analysis. An innovative diagnostic method for prostate and bladder cancer]. Urologe A. 2007;46(7):733-9.

248. Weissinger EM, Wittke S, Kaiser T, Haller H, Bartel S, Krebs R, et al. Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. Kidney Int. 2004;65(6):2426-34.

249. Theodorescu D, Schiffer E, Bauer HW, Douwes F, Eichhorn F, Polley R, et al. Discovery and validation of urinary biomarkers for prostate cancer. Proteomics Clin Appl. 2008;2(4):556-70.

250. Hanash S, Taguchi A, Wang H, Ostrin EJ. Deciphering the complexity of the cancer proteome for diagnostic applications. Expert Rev Mol Diagn. 2016;16(4):399-405.

251. Schiffer E, Mischak H, Zimmerli LU. Proteomics in gerontology: current applications and future aspects--a mini-review. Gerontology. 2009;55(2):123-37.

252. McKee JA, Kumar S, Ecelbarger CA, Fernandez-Llama P, Terris J, Knepper MA. Detection of Na(+) transporter proteins in urine. J Am Soc Nephrol. 2000;11(11):2128-32.

253. Franzen B, Auer G, Alaiya AA, Eriksson E, Uryu K, Hirano T, et al. Assessment of homogeneity in polypeptide expression in breast carcinomas shows widely variable expression in highly malignant tumors. Int J Cancer. 1996;69(5):408-14.

254. Hirano T, Fujioka K, Franzen B, Okuzawa K, Uryu K, Shibanuma H, et al. Relationship between TA01 and TA02 polypeptides associated with lung adenocarcinoma and histocytological features. Br J Cancer. 1997;75(7):978-85.

255. Paweletz CP, Ornstein DK, Roth MJ, Bichsel VE, Gillespie JW, Calvert VS, et al. Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma. Cancer Res. 2000;60(22):6293-7.

256. Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Res. 2002;62(13):3609-14.

257. Cheung PK, Woolcock B, Adomat H, Sutcliffe M, Bainbridge TC, Jones EC, et al. Protein profiling of microdissected prostate tissue links growth differentiation factor 15 to prostate carcinogenesis. Cancer Res. 2004;64(17):5929-33.

258. Check E. Proteomics and cancer: running before we can walk? Nature. 429. England2004. p. 496-7.

259. Okano T, Kondo T, Fujii K, Nishimura T, Takano T, Ohe Y, et al. Proteomic signature corresponding to the response to gefitinib (Iressa, ZD1839), an epidermal growth factor receptor tyrosine kinase inhibitor in lung adenocarcinoma. Clin Cancer Res. 2007;13(3):799-805.

260. Taguchi F, Solomon B, Gregorc V, Roder H, Gray R, Kasahara K, et al. Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. J Natl Cancer Inst. 2007;99(11):838-46.

261. de Wit M, Kant H, Piersma SR, Pham TV, Mongera S, van Berkel MPA, et al.

262. Marshall T, Williams K. Two-dimensional electrophoresis of human urinary proteins following concentration by dye precipitation. Electrophoresis. 1996;17(7):1265-72.

263. Ahlquist DA, Zou H, Domanico M, Mahoney DW, Yab TC, Taylor WR, et al. Next-generation stool DNA test accurately detects colorectal cancer and large adenomas. Gastroenterology. 2012;142(2):248-56; quiz e25-6.

264. Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, et al. Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. BMC Med. 2011;9:133.

265. Fung KY, Tabor B, Buckley MJ, Priebe IK, Purins L, Pompeia C, et al. Blood-based protein biomarker panel for the detection of colorectal cancer. PLoS One. 2015;10(3):e0120425.

266. Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A. 1979;76(9):4350-4.

267. Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis. 1995;16(6):1034-59.

268. Hochstrasser DF, Sanchez JC, Appel RD. Proteomics and its trends facing nature's complexity. Proteomics. 2002;2(7):807-12.

269. Wulfkuhle JD, Liotta LA, Petricoin EF. Proteomic applications for the early detection of cancer. Nat Rev Cancer. 2003;3(4):267-75.

270. Verma M, Wright GL, Jr., Hanash SM, Gopal-Srivastava R, Srivastava S. Proteomic approaches within the NCI early detection research network for the discovery and identification of cancer biomarkers. Ann N Y Acad Sci. 2001;945:103-15.

271. Fliser D, Novak J, Thongboonkerd V, Argiles A, Jankowski V, Girolami MA, et al. Advances in urinary proteome analysis and biomarker discovery. J Am Soc Nephrol. 2007;18(4):1057-71.

272. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet. 2002;359(9306):572-7.

273. Hampel DJ, Sansome C, Sha M, Brodsky S, Lawson WE, Goligorsky MS. Toward proteomics in uroscopy: urinary protein profiles after radiocontrast medium administration. J Am Soc Nephrol. 2001;12(5):1026-35.

274. Chen YD, Zheng S, Yu JK, Hu X. Artificial neural networks analysis of surface-enhanced laser desorption/ionization mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population. Clin Cancer Res. 2004;10(24):8380-5.

275. Yu JK, Chen YD, Zheng S. An integrated approach to the detection of colorectal cancer utilizing proteomics and bioinformatics. World J Gastroenterol. 2004;10(21):3127-31.

276. Cao P, Moini M. Analysis of peptides, proteins, protein digests, and whole human blood by capillary electrophoresis/electrospray ionization-mass spectrometry using an in-capillary electrode sheathless interface. J Am Soc Mass Spectrom. 1998;9(10):1081-8.

277. Dolnik V, Hutterer KM. Capillary electrophoresis of proteins 1999-2001. Electrophoresis. 2001;22(19):4163-78.

278. Frantzi M, Metzger J, Banks RE, Husi H, Klein J, Dakna M, et al. Discovery and validation of urinary biomarkers for detection of renal cell carcinoma. J Proteomics. 2014;98:44-58.

279. Frantzi M, van Kessel KE, Zwarthoff EC, Marquez M, Rava M, Malats N, et al. Development and Validation of Urine-based Peptide Biomarker Panels for Detecting Bladder Cancer in a Multi-center Study. Clin Cancer Res. 2016;22(16):4077-86.

280. Schonemeier B, Metzger J, Klein J, Husi H, Bremer B, Armbrecht N, et al. Urinary Peptide Analysis Differentiates Pancreatic Cancer From Chronic Pancreatitis. Pancreas. 2016;45(7):1018-26.

281. Belczacka I, Latosinska A, Siwy J, Metzger J, Merseburger AS, Mischak H, et al. Urinary CE-MS peptide marker pattern for detection of solid tumors. Sci Rep. 2018;8(1):5227.

282. National Cancer Registration and Analysis Service. Colorectal cancer 2016 [Available at: www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/colorectal_cancer/].

283. Thorne K, Hutchings HA, Elwyn G. The effects of the Two-Week Rule on NHS colorectal cancer diagnostic services: a systematic literature review. BMC Health Serv Res. 2006;6:43.

284. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-12.

285. National Institute for Health (National Heart Lung and Blood Institute). Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. [Available at: www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools].

286. Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36(3):1-48.

287. JJ M. The Inverse of the Freeman – Tukey Double Arcsine Transformation. The American Statistician. 1978;34(4):138.

288. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Bmj. 1997;315(7109):629-34.

289. Mozdiak E, Tsertsvadze A, McFarlane M, Widlak M, Tabuso M, Dunlop A, et al. The effect of the 2-week wait referral system on the detection of and mortality from colorectal cancer: protocol of a systematic review and meta-analysis. Syst Rev. 2016;5(1):182.

290. John SK, George S, Howell RD, Primrose JN, Fozard JB. Validation of the lower gastrointestinal electronic referral protocol. Br J Surg. 2008;95(4):506-14.

291. Borowski DW, Cawkwell S, Zaidi SM, Toward M, Maguire N, Gill TS. Primary care referral practice, variability and socio-economic deprivation in colorectal cancer. Colorectal Dis. 2016;18(11):1072-9.

292. Bevis PM, Donaldson OW, Card M, Durdey P, Thomas MG, Sylvester PA, et al. The association between referral source and stage of disease in patients with colorectal cancer. Colorectal Dis. 2008;10(1):58-62.

293. Schneider C, Bevis PM, Durdey P, Thomas MG, Sylvester PA, Longman RJ. The association between referral source and outcome in patients with colorectal cancer. Surgeon. 2013;11(3):141-6.

294. Hall SJ, Peacock JD, Cochrane LA, Peacock O, Tierney GM, Tou SI, et al. The bowel cancer awareness campaign 'Be Clear on Cancer': sustained increased pressure on

resources and over-accessed by higher social grades with no increase in cancer detected. Colorectal Dis. 2016;18(2):195-9.

295. Peacock O, Clayton S, Atkinson F, Tierney GM, Lund JN. 'Be Clear on Cancer': the impact of the UK National Bowel Cancer Awareness Campaign. Colorectal Dis. 2013;15(8):963-7.

296. Hammond TM, Fountain G, Cuthill V, Williams J, Porrett TR, Lunniss PJ. Straight to test. Results of a pilot study in a hospital serving an inner city population. Colorectal Dis. 2008;10(6):569-76.

297. Aslam MI, Chaudhri S, Singh B, Jameson JS. The "two-week wait" referral pathway is not associated with improved survival for patients with colorectal cancer. Int J Surg. 2017;43:181-5.

298. Bhangu A, Khan M, Roberts L, Reynolds A, Desai A, Mathew G. Detection and survival of colorectal cancer from a 2 week wait service. Surgeon. 2011;9(2):78-82.

299. Glancy DG, Card M, Sylvester PA, Thomas MG, Durdey P, Callaway M, et al. Fast-track barium enema: meeting the two-week wait rule for patients with suspected colorectal cancer. Colorectal Dis. 2005;7(3):241-4.

300. Khong TL, Naik K, Sivakumar R, Shah S. Impact of the United Kingdom national bowel cancer awareness campaigns 2012 on colorectal cancer services and patient survival. Colorectal Dis. 2015;17(12):1055-61.

301. Patel K, Doulias T, Hoad T, Lee C, Alberts JC. Primary-to-secondary care referral experience of suspected colorectal malignancy in young adults. Ann R Coll Surg Engl. 2016;98(5):308-13.

302. Saratzis A, Winter-Beatty J, El-Sayed C, Pande R, Harmston C. Colorectal cancer screening characteristics of patients presenting with symptoms of colorectal cancer and effect on clinical outcomes. Ann R Coll Surg Engl. 2015;97(5):369-74.

303. Currie AC, Evans J, Smith NJ, Brown G, Abulafi AM, Swift RI. The impact of the two-week wait referral pathway on rectal cancer survival. Colorectal Dis. 2012;14(7):848-53.

304. Neal RD, Allgar VL, Ali N, Leese B, Heywood P, Proctor G, et al. Stage, survival and delays in lung, colorectal, prostate and ovarian cancer: comparison between diagnostic routes. Br J Gen Pract. 2007;57(536):212-9.

305. Royle TJ, Ferguson HJ, Mak TW, Simpson JA, Thumbe V, Bhalerao S. Same-day assessment and management of urgent (2-week wait) colorectal referrals: an analysis of the outcome of 1606 patients attending an endoscopy unit-based colorectal clinic. Colorectal Dis. 2014;16(5):O176-81.

306. Aljarabah MM, Borley NR, Goodman AJ, Wheeler JM. Referral letters for 2-week wait suspected colorectal cancer do not allow a 'straight-to-test' pathway. Ann R Coll Surg Engl. 2009;91(2):106-9.

307. Allgar VL, Neal RD, Ali N, Leese B, Heywood P, Proctor G, et al. Urgent GP referrals for suspected lung, colorectal, prostate and ovarian cancer. Br J Gen Pract. 2006;56(526):355-62.

308. Banerjea A, Voll J, Chowdhury A, Siddika A, Thomson S, Briggs R, et al. Straight-to-test colonoscopy for 2-week-wait referrals improves time to diagnosis of colorectal cancer and is feasible in a high-volume unit. Colorectal Dis. 2017;19(9):819-26.

309. Barwick TW, Scott SB, Ambrose NS. The two week referral for colorectal cancer: a retrospective analysis. Colorectal Dis. 2004;6(2):85-91.

310. Beggs AD, Bhate RD, Irukulla S, Achiek M, Abulafi AM. Straight to colonoscopy: the ideal patient pathway for the 2-week suspected cancer referrals? Ann R Coll Surg Engl. 2011;93(2):114-9.

311. Couch DG, Murphy JH, Boyle KM, Hemingway DM. Straight to flexible sigmoidoscopy: rationalization of 2-week wait referrals in suspected colorectal cancer. Colorectal Dis. 2015;17(11):980-3.

312. Debnath D CR, Dielehner N, Gunning KA. Rapid access colorectal cancer referral: A significant association between cancer diagnosis and compliance with the guidelines. In: Ltd BS, editor. Annual Meeting of the Association of Coloproctology of Great Britain and Ireland2003. p. P023.

313. Eccersley AJ, Wilson EM, Makris A, Novell JR. Referral guidelines for colorectal cancer--do they work? Ann R Coll Surg Engl. 2003;85(2):107-10.

314. Gaunt A, Fisher RM, Robertson CS. Is CT scan a useful investigation for elderly patients with suspected colorectal cancer? . Association of coloproctology of Great Britain and Ireland Annual Meeting 2011; Birmingham, UK: Blackwell Publishing Ltd; 2011. p. Poster 16.

315. Khan KY, Xiong T, McCafferty I, Riley P, Ismail T, Lilford RJ, et al. Frequency and impact of extracolonic findings detected at computed tomographic colonography in a symptomatic population. Br J Surg. 2007;94(3):355-61.

316. Leung E, Grainger J, Bandla N, Wong L. The effectiveness of the '2-week wait' referral service for colorectal cancer. Int J Clin Pract. 2010;64(12):1671-4.

317. Maruthachalam K, Stoker E, Chaudhri S, Noblett S, Horgan AF. Evolution of the two-week rule pathway--direct access colonoscopy vs outpatient appointments: one year's experience and patient satisfaction survey. Colorectal Dis. 2005;7(5):480-5.

318. Mukherjee S, Fountain G, Stalker M, Williams J, Porrett TR, Lunniss PJ. The 'straight to test' initiative reduces both diagnostic and treatment waiting times for colorectal cancer: outcomes after 2 years. Colorectal Dis. 2010;12(10 Online):e250-4.

319. Padwick RT, Bajwa AA, Shaw A, Leung E, Francombe J, Stellakis ML. The Two-Week Referral System for colorectal cancer--not fit for purpose. Int J Colorectal Dis. 2013;28(11):1531-4.

320. Panagiotopoulou IG, Fitzrol D, Parker RA, Kuzhively J, Luscombe N, Wells AD, et al. The yield of colorectal cancer among fast track patients with normocytic and microcytic anaemia. Ann R Coll Surg Engl. 2014;96(4):289-93.

321. Patel RK, Sayers AE, Seedat S, Altayeb T, Hunter IA. The 2-week wait service: a UK tertiary colorectal centre's experience in the early identification of colorectal cancer. Eur J Gastroenterol Hepatol. 2014;26(12):1408-14.

322. Rai S, Kelly MJ. Prioritization of colorectal referrals: a review of the 2-week wait referral system. Colorectal Dis. 2007;9(3):195-202.

323. Rai S, Ballal M, Thomas WM, Miller AS, Jameson JS, Steward WP. Assessment of a patient consultation questionnaire-based scoring system for stratification of outpatient risk of colorectal cancer. Br J Surg. 2008;95(3):369-74.

324. Raje D, La Touche S, Mukhtar H, Oshowo A, Ingham Clark C. Changing trends in the management of colorectal cancers and its impact on cancer waiting times. Colorectal Dis. 2006;8(2):140-4.

325. Shaw AG, Lund JN, Longman C, Tierney GM, Goddard AF. The misuse of the faecal occult blood test under the lower gastrointestinal two week wait rule. Colorectal Dis. 2009;11(1):94-6.

326. Shaw AG, Simpson J, Tierney G, Goddard AF, Reynolds JR, Lund JN. Referral of patients with iron deficiency anaemia under the lower gastrointestinal two-week wait rule. Colorectal Dis. 2008;10(3):294-7.

327. Smith RA, Oshin O, McCallum J, Randles J, Kennedy S, Delamere S, et al. Outcomes in 2748 patients referred to a colorectal two-week rule clinic. Colorectal Dis. 2007;9(4):340-3.

328. Sorelli PG, Iliadis AD, Payne JG. The effectiveness of a rapid-access flexible sigmoidoscopy clinic in a district hospital. Int Surg. 2014;99(4):374-8.

329. Spencer H, Sakellariou V, Ullah S, Fulls E, Donnelly M. Gastrointestinal cancer and the 'two-week wait'. Clin Med (Lond). 2004;4(4):383.

330. Vaughan-Shaw PG, Cutting JE, Borley NR, Wheeler JM. Repeat 2-week wait referrals for colorectal cancer. Colorectal Dis. 2013;15(3):292-7.

331. Vulliamy P, McCluney S, Raouf S, Banerjee S. Trends in urgent referrals for suspected colorectal cancer: an increase in quantity, but not in quality. Ann R Coll Surg Engl. 2016;98(8):564-7.

332. Chohan DP, Goodwin K, Wilkinson S, Miller R, Hall NR. How has the 'two-week wait' rule affected the presentation of colorectal cancer? Colorectal Dis. 2005;7(5):450-3.

333. Courtney ED, Chong D, Tighe R, Easterbrook JR, Stebbings WS, Hernon J. Screen-detected colorectal cancers show improved cancer-specific survival when compared with cancers diagnosed via the 2-week suspected colorectal cancer referral guidelines. Colorectal Dis. 2013;15(2):177-82.

334. Dua RS, Brown VS, Loukogeorgakis SP, Kallis G, Meleagros L. The two-week rule in colorectal cancer. Can it deliver its promise? Int J Surg. 2009;7(6):521-5.

335. Flashman K, O'Leary DP, Senapati A, Thompson MR. The Department of Health's "two week standard" for bowel cancer: is it working? Gut. 2004;53(3):387-91.

336. Rao GN, Basnyat P, Taffinder N, Hudson C, Insall R. Reducing surgical outpatient waiting is not the solution to meeting the 2005 colorectal cancer target. Colorectal Dis. 2006;8(2):135-9.

337. Thornton L, Reader H, Stojkovic S, Allgar V, Woodcock N. Has the 'Fast-Track' referral system affected the route of presentation and/or clinical outcomes in patients with colorectal cancer? World J Surg Oncol. 2016;14(1):158.

338. Walsh S, Bruce C, Bennington S, Ravi S. The fourteen-day rule and colorectal cancer. Ann R Coll Surg Engl. 2002;84(6):386-8.

339. Bowel Cacncer UK. Staging and grading 2018 [Available at: www.bowelcanceruk.org.uk/about-bowel-cancer/diagnosis/staging-and-grading/].

340. Rees CJ TGS, Rutter MD, Baragwanath P, Pullan R, Feeney M, Haslam N. UK Key Performance Indicators and Quality Assurance Standards for Colonoscopy.: BSG/JAG.

341. National Cancer Registration and Analysis Service. Routes to Diagnosis - NCIN Data Briefing [Available at: www.ncin.org.uk/publications/data_briefing/routes_to_diagnosis].

342. National Cacner Research Insitute. Colorectal Cancer Survival by Stage 2009 [Available at: www.ncin.org.uk/databriefings].

343. National Cancer Registration and Analysis Service. National Report on Colorectal Staging Data Quality Report London, United Kingdom2009 [Available at: http://www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/colorectal_cancer/].

344. Hamilton W, Lancashire R, Sharp D, Peters TJ, Cheng KK, Marshall T. The importance of anaemia in diagnosing colorectal cancer: a case-control study using electronic primary care records. Br J Cancer. 2008;98(2):323-7.

345. John SKP, Jones OM, Horseman N, Thomas P, Howell RD, Fozard JBJ. Inter general practice variability in use of referral guidelines for colorectal cancer.

346. Zhou Y, Mendonca SC, Abel GA, Hamilton W, Walter FM, Johnson S, et al. Variation in 'fast-track' referrals for suspected cancer by patient characteristic and cancer diagnosis: evidence from 670 000 patients with cancers of 35 different sites. Br J Cancer. 2018;118(1):24-31.

347. Department of Health. Direct access to diagnostic tests for cancer: best practice referral pathways for general practitioners 2012 [Available at: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_133510].

348. Rabeneck L, Rumble RB, Thompson F, Mills M, Oleschuk C, Whibley A, et al. Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening. Can J Gastroenterol. 2012;26(3):131-47.

349. Halloran SP, Launoy G, Zappa M. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Faecal occult blood testing. Endoscopy. 2012;44 Suppl 3:Se65-87.

350. Widlak MM, Thomas CL, Thomas MG, Tomkins C, Smith S, O'Connell N, et al. Diagnostic accuracy of faecal biomarkers in detecting colorectal cancer and adenoma in symptomatic patients. Aliment Pharmacol Ther. 2017;45(2):354-63.

351. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med. 2014;370(14):1287-97.

352. Redwood DG, Asay ED, Blake ID, Sacco PE, Christensen CM, Sacco FD, et al. Stool DNA Testing for Screening Detection of Colorectal Neoplasia in Alaska Native People. Mayo Clin Proc. 2016;91(1):61-70.

353. Parris R . Lonestar and FAIMS CORE (PAD) 4:Introduction to FAIMS 2016 [Available at: https://support.owlstonenanotech.com/hc/en-us/articles/212367046-Introduction-to-FAIMS-].

354. Muffett R. Theory of Headspace Sampling 2016 [Available at: https://support.owlstonenanotech.com/hc/en-us/articles/212366966-Theory-of-Headspace-Sampling].

355. G.A.S. FlavourSpec - User Manual. Germany2017.

356. Wilson AD, Baietto M. Advances in electronic-nose technologies developed for biomedical applications. Sensors (Basel). 2011;11(1):1105-76.

357. Manolis A. The diagnostic potential of breath analysis. Clin Chem. 1983;29(1):5-15.

358. Esfahani S, Sagar NM, Kyrou I, Mozdiak E, O'Connell N, Nwokolo C, et al. Variation in Gas and Volatile Compound Emissions from Human Urine as It Ages, Measured by an Electronic Nose. Biosensors (Basel). 2016;6(1).

359. Probert CS, 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;18(3):337-43.

360. Jobu K, Sun C, Yoshioka S, Yokota J, Onogawa M, Kawada C, et al. Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer. Biol Pharm Bull. 2012;35(4):639-42.

361. Chawla NV BK, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence. 2002;16.

362. Phillips CM, Zatarain JR, Nicholls ME, Porter C, Widen SG, Thanki K, et al. Upregulation of Cystathionine-beta-Synthase in Colonic Epithelia Reprograms Metabolism and Promotes Carcinogenesis. Cancer Res. 2017;77(21):5741-54.

363. Shen XJ, Rawls JF, Randall T, Burcal L, Mpande CN, Jenkins N, et al. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. Gut Microbes. 2010;1(3):138-47.

364. Marchesi JR, Dutilh BE, Hall N, Peters WH, Roelofs R, Boleij A, et al. Towards the human colorectal cancer microbiome. PLoS One. 2011;6(5):e20447.

365. Dulal S, Keku TO. Gut microbiome and colorectal adenomas. Cancer J. 2014;20(3):225-31.

366. Barratt J, Topham P. Urine proteomics: the present and future of measuring urinary protein components in disease. Cmaj. 2007;177(4):361-8.

367. Metzger J, Chatzikyrkou C, Broecker V, Schiffer E, Jaensch L, Iphoefer A, et al. Diagnosis of subclinical and clinical acute T-cell-mediated rejection in renal transplant patients by urinary proteome analysis. Proteomics Clin Appl. 2011;5(5-6):322-33.

368. Lim LC, Lim YM. Proteome Heterogeneity in Colorectal Cancer. Proteomics. 2018;18(3-4).

369. Zhang C, Leng W, Sun C, Lu T, Chen Z, Men X, et al. Urine Proteome Profiling Predicts Lung Cancer from Control Cases and Other Tumors. EBioMedicine. 2018;30:120-8.

370. Wittke S, Fliser D, Haubitz M, Bartel S, Krebs R, Hausadel F, et al. Determination of peptides and proteins in human urine with capillary electrophoresis-mass spectrometry, a suitable tool for the establishment of new diagnostic markers. J Chromatogr A. 2003;1013(1-2):173-81.

371. Jantos-Siwy J, Schiffer E, Brand K, Schumann G, Rossing K, Delles C, et al. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. J Proteome Res. 2009;8(1):268-81.

372. Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, et al. Recommendations for biomarker identification and qualification in clinical proteomics. Sci Transl Med. 2010;2(46):46ps2.

373. DeLeno JM . Receiver operating characteristic laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty. Uncertainty modelling analysis, Second International Symposium.1993:[318-25 pp.].

374. The Human Protein Atlas. The Human Protein Atlas: MEP1A 2018 [Available at: www.proteinatlas.org].

375. Qureshi A, Pervez S. Allred scoring for ER reporting and it's impact in clearly distinguishing ER negative from ER positive breast cancers. J Pak Med Assoc. 2010;60(5):350-3.

376. Zhang Z, Wang Y, Zhang J, Zhong J, Yang R. COL1A1 promotes metastasis in colorectal cancer by regulating the WNT/PCP pathway. Mol Med Rep. 2018;17(4):5037-42.

377. Zou X, Feng B, Dong T, Yan G, Tan B, Shen H, et al. Up-regulation of type I collagen during tumorigenesis of colorectal cancer revealed by quantitative proteomic analysis. J Proteomics. 2013;94:473-85.

378. Wang XQ, Tang ZX, Yu D, Cui SJ, Jiang YH, Zhang Q, et al. Epithelial but not stromal expression of collagen alpha-1(III) is a diagnostic and prognostic indicator of colorectal carcinoma. Oncotarget. 2016;7(8):8823-38.

379. Turashvili G, Bouchal J, Baumforth K, Wei W, Dziechciarkova M, Ehrmann J, et al. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. BMC Cancer. 2007;7:55.

380. Tilman G, Mattiussi M, Brasseur F, van Baren N, Decottignies A. Human periostin gene expression in normal tissues, tumors and melanoma: evidences for periostin production by both stromal and melanoma cells. Mol Cancer. 2007;6:80.

381. Tian ZQ, Li ZH, Wen SW, Zhang YF, Li Y, Cheng JG, et al. Identification of Commonly Dysregulated Genes in Non-small-cell Lung Cancer by Integrated Analysis of Microarray Data and qRT-PCR Validation. Lung. 2015;193(4):583-92.

382. Li J, Ding Y, Li A. Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. World J Surg Oncol. 2016;14(1):297.

383. Chatterjee M, Mohapatra S, Ionan A, Bawa G, Ali-Fehmi R, Wang X, et al. Diagnostic markers of ovarian cancer by high-throughput antigen cloning and detection on arrays. Cancer Res. 2006;66(2):1181-90.

384. Bond JS, Matters GL, Banerjee S, Dusheck RE. Meprin metalloprotease expression and regulation in kidney, intestine, urinary tract infections and cancer. FEBS Lett. 2005;579(15):3317-22.

385. Sterchi EE, Naim HY, Lentze MJ, Hauri HP, Fransen JA. N-benzoyl-L-tyrosyl-p-aminobenzoic acid hydrolase: a metalloendopeptidase of the human intestinal microvillus membrane which degrades biologically active peptides. Arch Biochem Biophys. 1988;265(1):105-18.

386. Matters GL, Manni A, Bond JS. Inhibitors of polyamine biosynthesis decrease the expression of the metalloproteases meprin alpha and MMP-7 in hormone-independent human breast cancer cells. Clin Exp Metastasis. 2005;22(4):331-9.

387. Lottaz D, Maurer CA, Hahn D, Buchler MW, Sterchi EE. Nonpolarized secretion of human meprin alpha in colorectal cancer generates an increased proteolytic potential in the stroma. Cancer Res. 1999;59(5):1127-33.

388. Lottaz D, Maurer CA, Noel A, Blacher S, Huguenin M, Nievergelt A, et al. Enhanced activity of meprin-alpha, a pro-migratory and pro-angiogenic protease, in colorectal cancer. PLoS One. 2011;6(11):e26450.

389. Rosmann S, Hahn D, Lottaz D, Kruse MN, Stocker W, Sterchi EE. Activation of human meprin-alpha in a cell culture model of colorectal cancer is triggered by the plasminogen-activating system. J Biol Chem. 2002;277(43):40650-8.

390. Wang X, Chen J, Wang J, Yu F, Zhao S, Zhang Y, et al. Metalloproteases meprin-a (MEP1A) is a prognostic biomarker and promotes proliferation and invasion of colorectal cancer. BMC Cancer. 2016;16:383.

391. England N. Implementing a timed colorectal cancer diagnostic pathway. London; 2018.

392. Lawler M, Alsina D, Adams RA, Anderson AS, Brown G, Fearnhead NS, et al. Critical research gaps and recommendations to inform research prioritisation for more effective prevention and improved outcomes in colorectal cancer. Gut. 2018;67(1):179-93.

393. Zarkavelis G, Boussios S, Papadaki A, Katsanos KH, Christodoulou DK, Pentheroudakis G. Current and future biomarkers in colorectal cancer. Ann Gastroenterol. 2017;30(6):613-21.

394. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001;93(14):1054-61.

395. Isella C, Terrasi A, Bellomo SE, Petti C, Galatola G, Muratore A, et al. Stromal contribution to the colorectal cancer transcriptome. Nat Genet. 2015;47(4):312-9.

396. Calon A, Lonardo E, Berenguer-Llergo A, Espinet E, Hernando-Momblona X, Iglesias M, et al. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. Nat Genet. 2015;47(4):320-9.

397. Dunne PD, McArt DG, Bradley CA, O'Reilly PG, Barrett HL, Cummins R, et al. Challenging the Cancer Molecular Stratification Dogma: Intratumoral Heterogeneity Undermines Consensus Molecular Subtypes and Potential Diagnostic Value in Colorectal Cancer. Clin Cancer Res. 2016;22(16):4095-104.

398. McDonald PJ, Strachan JA, Digby J, Steele RJ, Fraser CG. Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. Clin Chem Lab Med. 2011;50(5):935-40.

399. Deutsch EW, Csordas A, Sun Z, Jarnuczak A, Perez-Riverol Y, Ternent T, et al. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res. 2017;45(D1):D1100-d6.

400. National Cancer Institute. Early Detection Research Network 2018 [Available at: https://edrn.nci.nih.gov].

401. Health. GAfGA. Enabling genomic data sharing for the benefit of human health. [Available at: https://genomicsandhealth.org/].

402. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer. 2005;5(11):845-56.

Appendix

Appendix 1. Search strategy for the systematic review (Chapter Three)

Search strategies for Pubmed MEDLINE

1. ((("Referral and Consultation"[Mesh])) AND "Neoplasms"[Mesh]) AND "Waiting Lists"[Mesh]

2. ((((fast track OR rapid OR urgent))) AND ("Referral and Consultation"[Mesh])) AND "Neoplasms"[Mesh]

3. ((((((outcomes OR survival OR diagnosis OR detection))) AND ((gastrointestinal OR colorectal OR oesophageal OR stomach OR duodenal OR bowel OR rectal OR CRC))) AND ((TWW OR 2WW OR two week wait OR rapid access OR fast track OR urgent referral)))) AND cancer

Appendix 2. Quality assessment tool for systematic review (Chapter Three)

12/11/2017 Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies - NHLBI, NIH

Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Appendix 3. Table of patient and Study characteristics for abstract data

CONFERENCE ABSTRACTS

STUDY ID, YEAR	Study characteristics	Population characteristics	Intervention groups	Outcomes
ADEOSUN, 2002	**Design:** Prospective	**% Women:** n/s	**Intervention group (n= 319):** All lower GI tww referrals within recruitment period	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		2. Cancer stage at diagnosis
	No. of centres: 1	**Recruitment dates:** 1 year (date not stated)		3. Other diagnoses made
	Length of follow-up: n/s	**Total Total sample size:** 319		4. Adherence to targets
		Total drop-out/excluded: n/s		
ARYAL, 2003	**Design:** Prospective	**% Women:** 47	**Intervention group (n=50):** All newly diagnoses CRC within time period	1. Adherence to targets
	Study setting: Secondary care	**Median age (range):** mean 70 (34-97)		
	No. of centres: 1	**Recruitment dates:** April 2001-July 2002	**Comparator group (n=41):** Routine and soon referrals	

	Length of follow-up: n/s	**Total Total sample size:** 91		
		Total drop-out/excluded: n/s		
BROWNING, 2016	**Design:** Retrospective	**% Women: n/s**	**Intervention group (n=624):** All lower GI tww referrals within the recruitment period	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		2. Adherence to targets
	No. of centres: 1	**Recruitment dates:** March-October 2015		
	Length of follow-up: n/s	**Total Total sample size:** 624		
		Total drop-out/excluded: n/s		
CHODHARY, 2005	**Design:** Not stated	**% Women:** n/s	**Intervention group (n= 931):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** Aug 2000-Jul 2003		
	Length of follow-up: n/s	**Total sample size:** 931		

		Total drop-out/excluded: n/s		
FLASHMAN, 2001	**Design:** Not stated	**% Women:** n/s	**Intervention group (n= 284):** All lower GI referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** 1st Jul 2000-31st Oct 2000		
	Length of follow-up: n/s	**Total sample size:** 884	**Comparator group (n= 600):** Routine lower GI referrals	
		Total drop-out/excluded: n/s		
GANDY, 2002	**Design:** Prospective observational study	**% Women:** n/s	**Intervention group (n= 543):** All lower GI referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		2. Adherence to targets
	No. of centres: 1	**Recruitment dates:** Jun 2000-Nov 2001	**Comparator group (n=1961):** Non- tww lower GI referrals	
	Length of follow-up: n/s	**Total sample size:** 2504		
		Total drop-out/excluded: n/s		
HERIOT, 2001	**Design:** Prospective	**% Women:** n/s		1. Cancer conversion rate

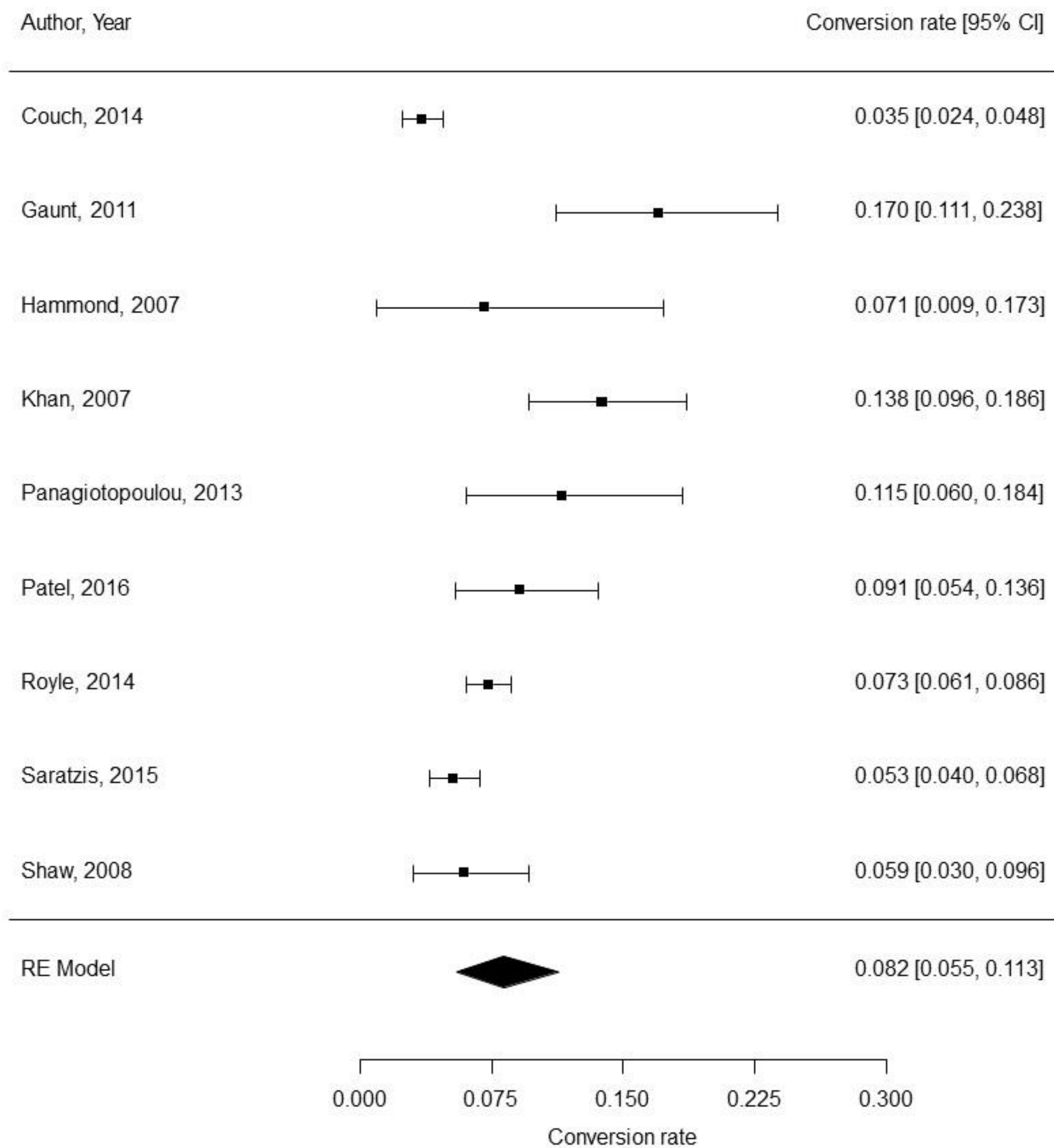
	Study setting: Secondary care	**Median age (range):** n/s	**Intervention group (n= 37):** All lower GI tww referrals within time frame	
	No. of centres: 1	**Recruitment dates:** 3 months (directly after tww implementation)		
	Length of follow-up: n/s	**Total sample size:** 37		
		Total drop-out/excluded: n/s		
MACDONALD, 2002	**Design:** Prospective	**% Women:** n/s	**Intervention group (N= 50):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** Not stated		
	Length of follow-up: n/s	**Total sample size:** 50		
		Total drop-out/excluded: n/s		
MARTIN, 2016	**Design:** Prospective	**% Women:** n/s	**Intervention group (N= 568):** All lower GI tww referrals within time frame	1. Adherence to targets
	Study setting: Secondary care	**Median age (range):** n/s		

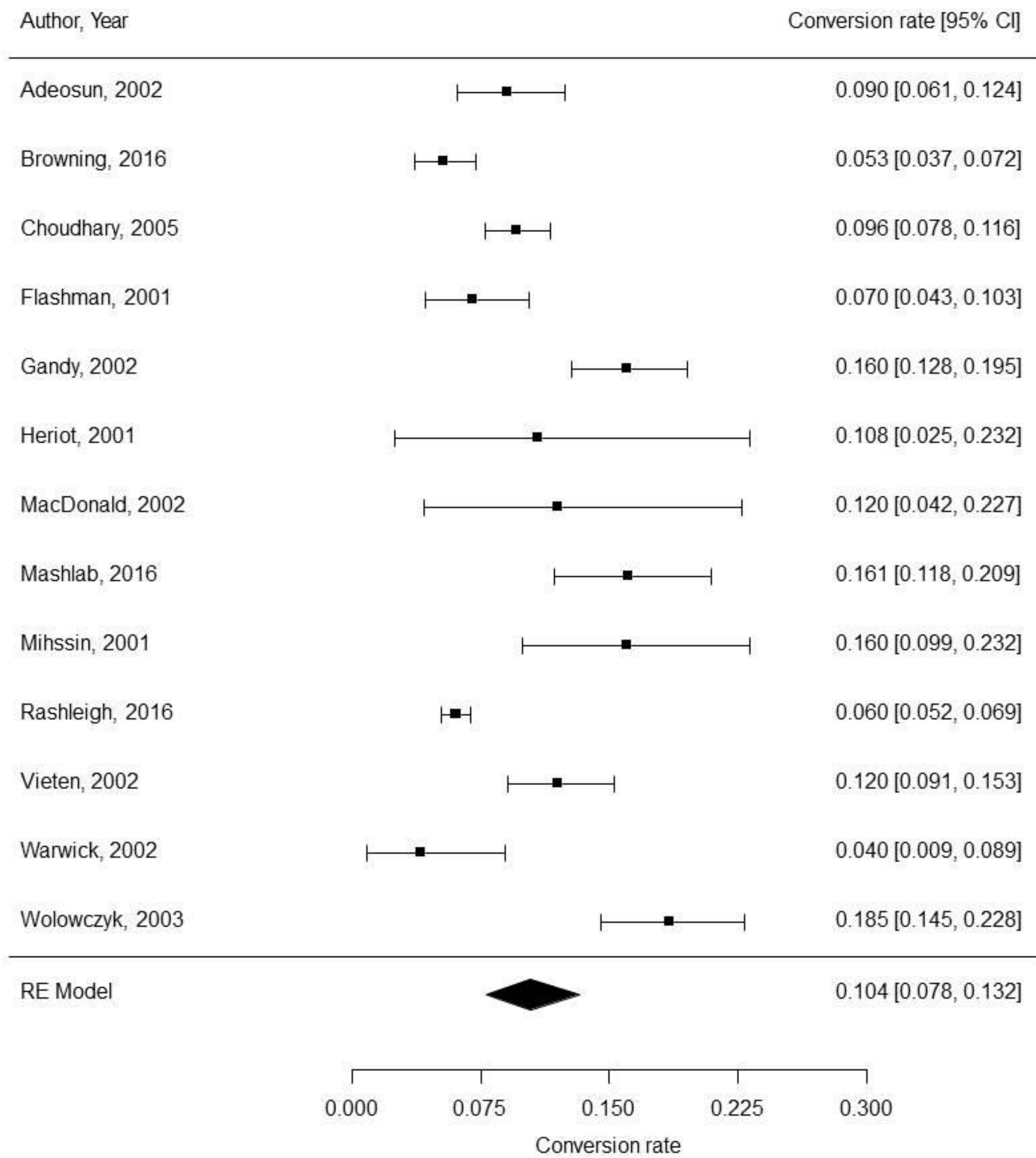
	No. of centres: 1	**Recruitment dates:** 3 months (not stated which)		
	Length of follow-up: n/s	**Total sample size:** 568		
		Total drop-out/excluded: n/s		
MASHLAD, 2016	**Design:** Retrospective	**% Women:** n/s	**Intervention group (N= 1000):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** 01/12/14-13/7/2015		
	Length of follow-up: n/s	**Total sample size:** 1000		
		Total drop-out/excluded: n/s		
MIHSSIN, 2001	**Design:** Prospective	**% Women:** n/s	**Intervention group (n= 120):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		2. Adherence to targets
	No. of centres: 1	**Recruitment dates:** 1st Jul 2000-31st Oct 2000		

	Length of follow-up: n/s	**Total sample size:** 120		
		Total drop-out/excluded: n/s		
RASHLEIGH, 2016	**Design:** Retrospective	**% Women:** n/s	**Intervention group (n= 2994):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** 2011-2014		
	Length of follow-up: n/s	**Total sample size:** 2994		
		Total drop-out/excluded: n/s		
VIETEN, 2002	**Design:** Prospective	**% Women:** n/s	**Intervention group (n= 420)**: All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** Mar 2000-Dec 2001		
	Length of follow-up: n/s	**Total sample size:** 420		

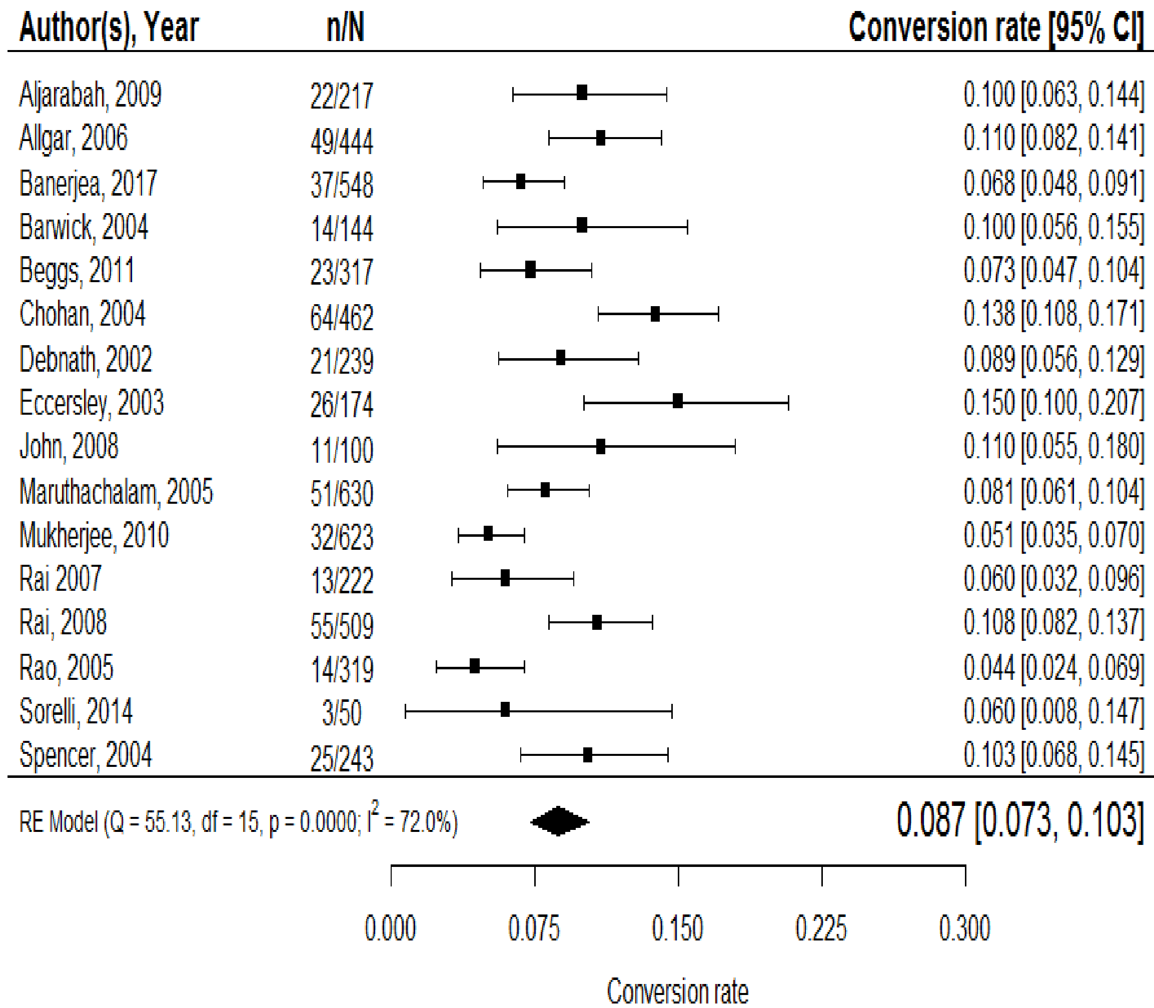
Study	Design/setting	Participants	Interventions	Outcomes
		Total drop-out/excluded: n/s		
WARWICK, 2002	**Design:** Audit	**% Women:** n/s	**Intervention group (n=100):** All lower GI tww referrals within time frame	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** (32-82)		
	No. of centres: 1	**Recruitment dates:** Not stated		
	Length of follow-up: n/s	**Total sample size:** 100		
		Total drop-out/excluded: n/s		
WOLOWCZYK, 2003	**Design:** Audit	**% Women:** n/s	**Intervention group (n= 335):** All lower GI tww referrals within the three recruitment periods	1. Cancer conversion rate
	Study setting: Secondary care	**Median age (range):** n/s		
	No. of centres: 1	**Recruitment dates:** Three 3-month periods in each year from 2000-2002		
	Length of follow-up: n/s	**Total sample size:** 342		
		Total drop-out/excluded: 7		

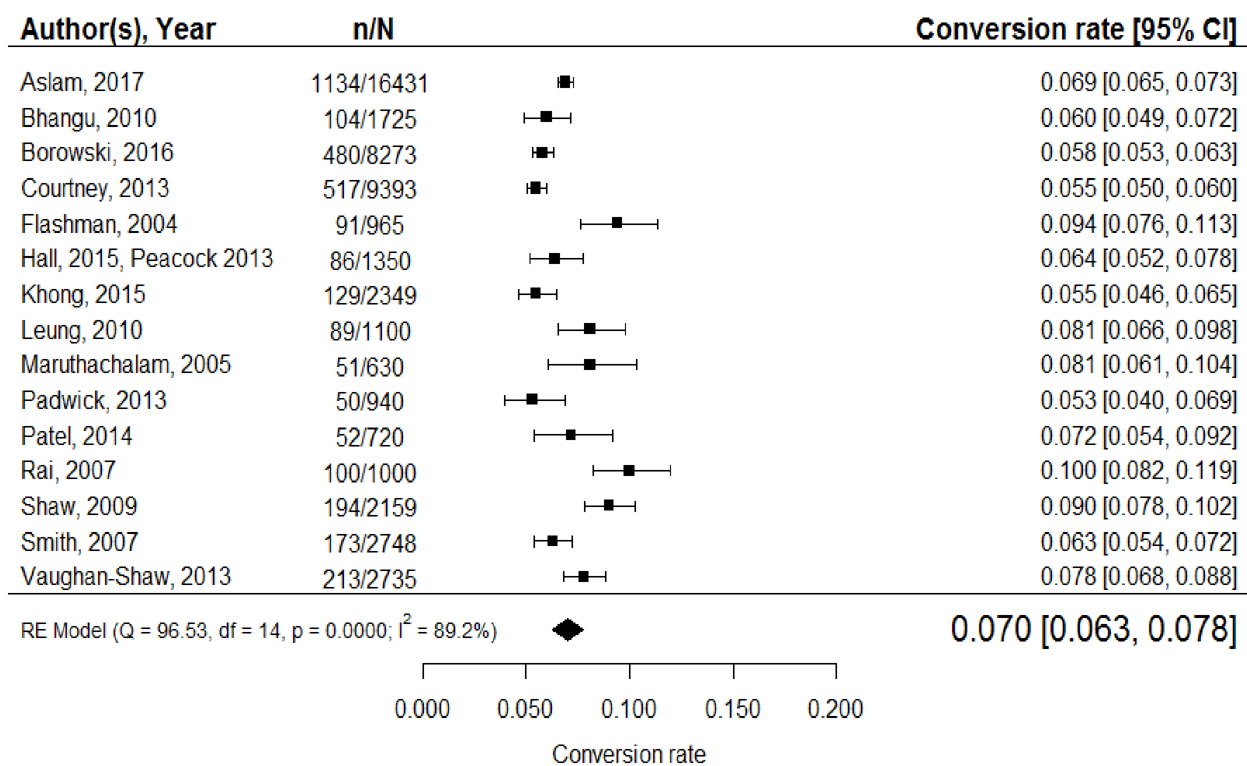
Appendix 4. Forest plot for cancer conversion rate in the subset (4a) and abstract (4b) groups.




Author, Year
Conversion rate [95% CI]
Adeosun, 2002
0.090 [0.061, 0.124]
Browning, 2016
0.053 [0.037, 0.072]
Choudhary, 2005
0.096 [0.078, 0.116]
Flashman, 2001
0.070 [0.043, 0.103]
Gandy, 2002
0.160 [0.128, 0.195]
Heriot, 2001
0.108 [0.025, 0.232]
MacDonald, 2002
0.120 [0.042, 0.227]
Mashlab, 2016
0.161 [0.118, 0.209]
Mihssin, 2001
0.160 [0.099, 0.232]
Rashleigh, 2016
0.060 [0.052, 0.069]
Vieten, 2002
0.120 [0.091, 0.153]
Warwick, 2002
0.040 [0.009, 0.089]
Wolowczyk, 2003
0.185 [0.145, 0.228]
RE Model
0.104 [0.078, 0.132]
0.000
0.075
0.150
0.225
0.300
Conversion rate

Appendix 5. Forest plot of cancer conversion rate for full publications split into low (top figure) and high (low figure) study population size.




Author(s), Year
n/N
Conversion rate [95% CI]
Aslam, 2017
1134/16431
0.069 [0.065, 0.073]
Bhangu, 2010
104/1725
0.060 [0.049, 0.072]
Borowski, 2016
480/8273
0.058 [0.053, 0.063]
Courtney, 2013
517/9393
0.055 [0.050, 0.060]
Flashman, 2004
91/965
0.094 [0.076, 0.113]
Hall, 2015, Peacock 2013
86/1350
0.064 [0.052, 0.078]
Khong, 2015
129/2349
0.055 [0.046, 0.065]
Leung, 2010
89/1100
0.081 [0.066, 0.098]
Maruthachalam, 2005
51/630
0.081 [0.061, 0.104]
Padwick, 2013
50/940
0.053 [0.040, 0.069]
Patel, 2014
52/720
0.072 [0.054, 0.092]
Rai, 2007
100/1000
0.100 [0.082, 0.119]
Shaw, 2009
194/2159
0.090 [0.078, 0.102]
Smith, 2007
173/2748
0.063 [0.054, 0.072]
Vaughan-Shaw, 2013
213/2735
0.078 [0.068, 0.088]
RE Model (Q = 96.53, df = 14, p = 0.0000; I2 = 89.2%)
0.070 [0.063, 0.078]
0.000
0.050
0.100
0.150
0.200
Conversion rate