IDENTIFYING RISK FACTORS AND BIOMARKERS FOR COLORECTAL CANCER SCREENING

A thesis submitted for the degree of Doctor of Philosophy
at

University College London (UCL)

Darren S. Thomas

2019

Under the guidance of

Professor Usha Menon

&

Dr Aleksandra Gentry-Maharaj

I, Darren S. Thomas, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

'Declare the past, diagnose the present, foretell the future.'

Hippocrates

"In our society we have no major crimes…but we do have a detention camp full of would-be criminals."

Philip K. Dick

The Minority Report.

A Predilection for Prediction

Amidst the hullaballoo of a shanty, post-war Kampala, an intrepid surgeon-cum-researcher was baffled by something rather odd. Why, despite living 'largely off the land' and having little access to a most basic Ugandan healthcare, a puzzled Dennis Burkitt pondered, were the humble locals seldom stricken with the same bowel cancers he had seen afflict the clinics at the forefront of mid-twentieth-century medicine?

"Western diets are so low on bulk [fibre] and so dense in calories," an eagle-eyed Burkitt would later conclude, "that our intestines just don't pass enough volume to remain healthy."

Overindulging on growth signals, absconding arrest from the immune system, and perhaps even crossing borders to far-flung organs, bowel cancers tend to remain well-hidden for two-to-three decades. They may be there, squatting in the dank alimentary alleyway—silent, multiplying, biding their time—but before long their loitering will cause discomfort, sending an uneased patient for medical attention. Most often it is too late.

We know that the earlier we can detect these clumps of cancerous cells, the easier it is to carve them out. But leave just one cancerous cell behind, and one can split into two into four into *ad finitum*.

What's more, should some cells splinter from the mass and lodge themselves into distant organs, the challenge to remove every cancerous cell is too often insurmountable. Their early detection is key.

The National Health Service has since 2010 offered hope for better outcomes for the forty-odd thousand Britons afflicted with bowel cancer each year. By looking for smidgens of blood not visible to the naked eye in faeces—a major symptom of bowel cancer—it aims to sniff and then snuff out these cancers-in-hiding far earlier than normal. But while this screening programme is no doubt lifesaving, my research aims to make it even better—and get more bang for British taxpayer's bucks.

Burkitt's work in East Africa was groundbreaking. Yet, the many studies conducted by countless scientists since have unravelled the causes of bowel cancer to be more complex than his 1970s bestseller *Don't Forget Your Fibre* could envisage. But he *was* on to something.

Bowel cancers are still to this day a rarity in rural Africa and a major burden for developed countries. Coincidentally, those emigrating from countries of low risk soon acquire the risk of their new compatriots. The exuberances of modern living are unquestionably to blame.

Fags, booze, meaty menus, and our tendency to whittle away hours in front of the box all increase your risk to varying degrees. What's more, the maladies of modern living, like diabetes and inflammatory bowel diseases, can increase one's risk too.

My research will scrutinise how the demographics of women who develop bowel cancer in the UK vary from their cancer-free peers. How, for example, do they differ in their age, ethnicity, medical history, and lifestyles? And how much more likely is, say, a smoker or a drinker to develop bowel cancer to abstainers? A diabetic from a non-diabetic? And a sixty-year-old from a fifty-something?

I am toiling toward an algorithm which, when you plug in data about you and your lifestyle choices, crunches the numbers and spits out a personalised risk score. We could then tailor bowel screening based on your personalised risk score: the higher your risk, the more you should be kept an eye on.

But why does this matter? Invitees to screening are an exclusive club reserved for the 60–74-year olds—an age when bowel cancers are more likely to develop. But the issue lies in the math. Some of this group will be of high risk (and will benefit from frequent screening), while others will have a low personal risk (and will not benefit from frequent screening).

A tailored screening programme, bespoke for each person, would be more efficient. It would also be indiscriminate to age, throwing a lifeline to more of the 16% of all patients who receive their diagnosis before their screening invite.

As a bookish youngster, I thumbed my way through stacks of dog-eared paperbacks during the solitary Suffolk summers. I spent hours with Philip K. Dick's *Minority Report*. In some forty-odd pages, Dick depicts a fictional 'utopia' wherein the city's Precrime system would apprehend those 'accused not of crimes they have committed, but of crimes they will commit'. Wrongdoings were all but eradicated.

It was only later, undertaking this research—a stone's throw from Burkitt's 'retirement' laboratory in London—that this literary fanaticism took on a higher level of poignancy: I was vying to become Commissioner Anderton. What if, just as Anderton's invention would remedy events before they occurred, we could—to paraphrase Phillip K. Dick—treat bowel cancers accused not of symptoms they have shown but of symptoms they will show?

Shortlisted for the Medical Research Council
Max Perutz Science Writing Award 2017

Acknowledgments

There are many people I wish to thank, though I am sure words alone cannot express my gratitude. To my supervisors, Professor Usha Menon and Dr Aleksandra Gentry-Maharaj, thank you for your unfailing support, inspiration, and guidance. You have made the journey meaningful enough to justify its many challenges and for this I will forever be grateful. To Dr Andy Ryan, Dr Sophia Apostolidou, Dr Matthew Burnell, and Mr Richard Gunu for lending me their time, expertise, and advice in abundance. To the Gynaecological Cancer Research Centre for the always warm and stimulating environment I was fortunate to work within. I also wish to thank Dr Sioban Sen Gupta and Dr Angela Poulter for navigating me through the unusual and difficult times.

This research was the brainchild of Dr John Timms and Dr Wendy Alderton, to whom I am grateful for gambling on my admission to UCL and for their guidance during the early years. I am also indebted to the Medical Research Council, Abcodia, and to Dr Julie Barnes for their support and encouragement throughout.

Above all, this research would not have been possible without the women who took part in the UKCTOCS and for the many clinical, research, and administrative staff who brought it to fruition. Finally, the current work has been honed in response to the reviews of many unsung peers—thank you.

November 2018.

This thesis has been bettered on the thorough review and examination by Professor Robert Steele and Professor Nora Pashayan—thank you for your time and consideration.

May 2019.

Abbreviations & Acronyms

The following abbreviations and acronyms are used throughout and are best listed here:

BCSP	Bowel Cancer Screening Programme
BMI	Body Mass Index
CA125	Ovarian carcinoma antigen 125
CEA	Carcinoembryonic Antigen
CI	Confidence Interval
CR	Cancer Registry
CRC	Colorectal Cancer
CRCQ	Colorectal Cancer Questionnaire
CV	Coefficient of Variation
CYFRA21-1	Cytokeratin-19 fragment
DC	Death Certificate
DCD	Dermcidin
EHR	Electronic Health Record
FAP	Fibroblast Activation Protein
FIT	Faecal Immunochemical Test
FN	False Negative
FP	False Positive
FUQ	Follow-up Questionnaire
gFOBT	Guaiac Faecal Occult Blood Test
HES	Hospital Episode Statistics
HR	Hazard Ratio
ICD	International Classification of Diseases
IMD	Index of Multiple Deprivation
IQR	Interquartile Range
LC	Liquid Chromatography
MS/MS	Tandem Mass Spectrometry
NCIN	National Cancer Intelligence Network
NGRS	Non-genetic Risk Score
NPV	Negative Predictive Value
NR	Non-responder
OR	Odds Ratio
PEDW	Patient Episode Database for Wales
PPV	Positive Predictive Value
RQ	Recruitment Questionnaire
RR	Relative Risk
S100A8	Protein S100-A8
SIR	Standardised Incidence Ratio
SR	Self-reporting
TIMP-1	Metalloproteinase inhibitor 1
TMT	Tandem Mass Tag
TN	True Negative
TP	True Positive
UKCTOCS	UK Collaborative Trial of Ovarian Cancer Screening
VEGFA	Vascular Endothelial Growth Factor A

Abstract

Background. An inefficient selection of screenees and low uptake to stool-based screening undermines the potential benefits of the BCSP. The present thesis, through a series of cohort and case-control studies nested within the UKCTOCS involving 202,638 postmenopausal women, studies the viability of risk-stratified and blood-based screening for detecting asymptomatic CRCs.

Chapter 3. CRCs diagnosed during 2001–11 following randomisation to the UKCTOCS were identified via EHRs and self-report and later verified with their treating clinician. Notifications for CRC diagnoses were received for 1,085 women by May 2011, in which 514 has CRC, 24 had benign polyp, and 103 had neither diagnosis. Cancer & death registrations and Hospital Episode Statistics used in adjunct had a sensitivity of 98% and a positive predictive value of 92%. EHRs together are a reliable means for ascertaining events of CRC.

Chapter 4. A NGRS for risk-stratified screening was derived from a Cox proportional-hazards model of 202,323 women (4,134 CRCs) and 16 shortlisted covariates recorded in EHRs. Reclassifying the risk of a bootstrap resample (n 138,900 (2,678 CRCs)), 5.6% more potentially harmful screens can be undertaken to potentially detect the same proportion of cancers than criterium based on age alone (50–74 years). The top decile had a three-fold increased hazards over the geometric mean.

Chapter 5. CA125, CEA, CYFRA21-1, FAP, TIMP-1, and VEGFA were shortlisted from the literature and later evaluated as screening markers on immunoassay measurements of 386 longitudinal preclinical sera drawn 0–4 years before the diagnosis of benign adenoma and CRC and matched controls who remained cancer-free. CEA is elevated up to a year and four years in advance of diagnosis in 23 and 12% of cases, respectively. Other markers had sensitivities of 5–12 and 4–10% at the same lead times.

Chapter 6. Novel markers were sought through analyses of pools of preclinical sera using LC-MS/MS with quantitative Tandem Mass Tags. DCD and S100A8 were shortlisted on their differential and longitudinal expression in CRCs and controls and were evaluated using individual measurements obtained with immunoassays. DCD and S100A8 were elevated up to one year before diagnosis for only 10% of CRCs.

Chapter 7. The thesis concludes with an exploration of all multi-marker permutations. The top-ranked—a panel of CA125, CEA, FAP, TIMP-1 and VEGFA— had, at 95% specificity, a sensitivity of 23% for detecting CRCs up to one-year and 10% up to four years before diagnosis. The panel performed no better than CEA alone.

Afterword. The use of EHRs to ascertain events of CRC in BigData is justified. The NGRS requires refinements before informing risk-stratified screening. CEA is elevated up to one year before diagnosis in 23% of future CRC patients. No circulating proteins are viable as standalone tools for screening for CRC.

Impact Statement

The current work:

—lends evidence to support the use of routine Electronic Health Records for research on colorectal cancer.

—has verified and obtained data for 515 colorectal cancers diagnosed in women since their involvement in the UKCTOCS, which allows for the use of their donated blood for future research.

—provides a proof-of-concept on how Electronic Health Records may inform eligibility to the NHS Bowel Cancer Screening Programme.

—reports, many for the first time, the behaviour of eight proteins over four-years before the diagnoses of colorectal cancer.

—has reached a readership of non-specialists and laypersons through the Medical Research Council Max Perutz Science Writing Award.

List of Contents

Chapter 1. Preface

CRCs develop through a lengthy, premalignant phase when simple polypectomy provides a means of secondary prevention of malignant disease [1] and those that do become cancerous, but have not evolved invasive and metastatic traits, are highly treatable and have improved survival outcomes [2]. It is on these observations that screening the asymptomatic for CRC is an efficacious and cost-effective means for reducing related mortality [3,4]. Nested within the UK Collaborative Trial of Ovarian Cancer Screening involving 202,638 postmenopausal women, the present thesis works to address the shortcomings of the current BCSP, first implemented by the NHS in 2006.

1.1 Raison D'etre

The reason for being is owed to two suboptimal realities of the BCSP: I) that eligibility based on age alone is somewhat inefficient and II) uptake to the stool-based screening programme is notoriously low.

On the inefficient selection criteria, the BCSP partitions risk of CRC purely on age; assuming those aged 60–74 would benefit equally from fixed, biennial screening. (During the drafting of the current work, the UK National Screening Committee announced that screening would commence ten-years earlier—at age fifty [5]). In truth, CRC can develop below this arbitration—as 16% did in 2012 (data provided by the national cancer registries on request, February 2016). On the other hand, 0.5% (5,198 of 1,079,293) of screening-test findings were falsely positive during 2006–08, and were therefore potentially harmful on the undue anxiety and distress caused [6] or from the unnecessary complications (bowel perforation and/or haemorrhaging) that arise in 0.1–1% of diagnostic colonoscopies [7,8]. Albeit the harms of screening are a rarity, the frequency of which the population, perhaps unnecessarily, attend routine screening amounts to a marked number of persons potentially incurring harm. Given that a sufficient component of CRC susceptibility is non-genetic [9] and many risk factors—comorbidities, medications, anthropometrics, etc.—are recorded by EHRs during routine healthcare; could eligibility based on a NGRS be more efficient than the current criterium on age alone?

Uptake to the BCSP, meanwhile, is particularly low and this undermines the potential benefits CRC screening can provide. Just 54% of those eligible during 2006–08 returned a completed test kit [6], for example, and this rate is somewhat consistent with the proportion who partook in the Nottingham-based randomised-controlled trial (60%)

[10] and demonstration pilot (57%) in England and Scotland [11]. Recent evidence, too, suggests that uptake has decreased year-on-year since 2010 [12], while inequalities in uptake due to deprivation, ethnicity, and gender are widely reported [12–14] . There is at least some evidence that those of average-risk and with a history of refusing conventional screening prefer to undergo a blood test [15,16]. An alternative screening test would need to be of at least the same, if not greater, performance to the current gFOBT, which has a sensitivity of 13–50% for detecting asymptomatic CRC in an average-risk demographic, dependent on the lead time afforded [17,18]. Since commencing this research, the UK National Screening Committee advocated in 2015 for the transition to the next-generation occult blood test, the FIT, which is/was due to be implemented during 2018 [19]. This recommendation was based on evidence that use of the FIT over the gFOBT in the BCSP improved uptake, particularly in previous non-responders [20]. An efficacious blood-based screening test could, nonetheless, be used as a second-line alternative for those who refuse stool-based tests.

1.2 UKCTOCS

All studies to be presented are nested within the UK Collaborative Trial of Ovarian Screening and studies those who granted us permission for use of their collected samples and data in secondary, ethically approved studies. The UKCTOCS [ClinicalTrials.gov: NCT00058032] is/was a multi-centre randomised-controlled trial that aimed to inform on the effect of population-based ovarian cancer screening on related mortality in the UK [21,22]. During 2001–05, 202,638 postmenopausal women aged 50–74 years were recruited from primary care trusts across England, Northern Ireland, and Wales and randomised 1:1:2 to annual multimodal screening (n 50,640) via serum CA125 interpreted with the Risk of Ovarian Cancer Algorithm [23,24], annual transvaginal ultrasound screening (n 50,639), or no screening (n 101,359). Women received 7–11 screens until cessation on 31 December 2011, for a total of 345,570 multimodal and 327,775 ultrasound screening episodes. The UKCTOCS biobank archives all sera obtained from all 202,638 women at recruitment and all subsequent multimodal screening episodes.

1.3 Outline of Thesis

The present thesis consists of seven further chapters. The proceeding chapter is devoted to fleshing out the concepts introduced here in the preface; the burden of CRC, on how the natural history and clinical management of CRC lend itself to be an attractive

candidate for screening and concludes with a narrative of CRC screening in the UK from its initial trials and pilots through to the present day BCSP. In Chapter 3, diagnoses of CRC in the UKCTOCS cohort are identified through EHRs, self-report by postal questionnaire, and via UKCTCOS trial resources, and are thereafter verified with the woman's treating clinician. Not only does this enable the use of their archived serum samples for study in future Chapters, but also performance estimates for the use of EHRs and/or self-reporting for ascertaining CRC in future epidemiology studies can be estimated. It is here that the current performance estimates for cancer registries are revised [25,26], while for the first time estimates on the use of Hospital Episode Statistics and self-reporting for monitoring for colorectal cancer events are reported. Thereafter, and informed by the findings of the previous chapter, the viability of risk-stratified screening based on non-genetic variables recorded by EHRs is studied. In particular, can risk-stratified screening allocate the finite endoscopic resources more efficiently than the current eligibility based on age alone? The subsequent three chapters shall attempt to identify potential candidates for a blood-based screening test. Firstly, in Chapter 5, promising candidates are shortlisted from the literature and then evaluated on their ability to detect asymptomatic CRCs, using serial preclinical samples donated before the diagnosis of disease. Thereon, Chapter 6 attempts to identify novel markers in the laboratory using reputed proteomic methods. This study is one of few that have undertaken marker discovery work on preclinical samples. The last of the experimental chapters, Chapter 7 explores whether the screening performance of an individual marker can be complemented by one or more additional markers. The thesis concludes with summary of the work done in context of the literature, and it is here the future of CRC screening is speculated on.

1.4 Ethical Approval

Approval for the studies herein were granted by the National Research Ethics Service Committee of East Midlands, Derby [REC 13/EM/0191]. The UKCTOCS was formerly approved by the UK North West Medical Research Ethics Service [MREC 00/8/34].

1.5 Funding

The present study was supported by the Medical Research Council Industrial CASE Studentship [MR/J006718/1]. The Medical Research Council had no involvement in the design, undertaking, or reporting of all studies. The UKCTOCS was funded by the

Medical Research Council, Cancer Research UK, the Department of Health, and The Eve Appeal.

1.6 References

[1] S.J. Winawer, A.G. Zauber, M.N. Ho, M.J. O'Brien, L.S. Gottlieb, S.S. Sternberg, J.D. Waye, M. Schapiro, J.H. Bond, J.F. Panish, Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup, N. Engl. J. Med. 329 (1993) 1977–81. doi:10.1056/NEJM199312303292701.

[2] Office for National Statistics, Cancer survival by stage at diagnosis for England (experimental statistics): adults diagnosed 2012, 2013 and 2014 and followed up to 2015, 2016.

[3] P. Hewitson, P. Glasziou, E. Watson, B. Towler, L. Irwig, Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update, Am J Gastroenterol. 103 (2008) 1541–1549. doi:10.1111/j.1572-0241.2008.01875.x.

[4] P. Tappenden, J. Chilcott, S. Eggington, H. Sakai, J. Karnon, J. Patnick, Option appraisal of population-based colorectal cancer screening programmes in England, Gut. 56 (2007) 677–684. doi:10.1136/gut.2006.095109.

[5] UK National Screening Committee, Note of the meeting held on the 29 June 2018 in Edinburgh, Scotland, 2018.

[6] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut. 61 (2012) 1439–46. doi:10.1136/gutjnl-2011-300843.

[7] F.A. Macrae, K.G. Tan, C.B. Williams, Towards safer colonoscopy: a report on the complications of 5000 diagnostic or therapeutic colonoscopies, Gut. 24 (1983) 376–83.

[8] M.H. Robinson, J.D. Hardcastle, S.M. Moss, S.S. Amar, J.O. Chamberlain, N.C. Armitage, J.H. Scholefield, C.M. Mangham, The risks of screening: data from the Nottingham randomised controlled trial of faecal occult blood screening for colorectal cancer, Gut. 45 (1999) 588–92.

[9] P. Lichtenstein, N. V Holm, P.K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, K. Hemminki, Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland, N Engl J Med. 343 (2000) 78–85. doi:10.1056/NEJM200007133430201.

[10] J.D. Hardcastle, J.O. Chamberlain, M.H. Robinson, S.M. Moss, S.S. Amar, T.W. Balfour, P.D. James, C.M. Mangham, Randomised controlled trial of faecal-occult-blood screening for colorectal cancer, Lancet. 348 (1996) 1472–1477. doi:10.1016/s0140-6736(96)03386-7.

[11] UK Colorectal Cancer Screening Pilot Group, Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom, BMJ. 329 (2004) 133. doi:10.1136/bmj.38153.491887.7C.

[12] Y. Hirst, S. Stoffel, G. Baio, L. McGregor, C. von Wagner, Uptake of the English Bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out, Eur. J. Cancer. 103 (2018) 267–273. doi:10.1016/j.ejca.2018.07.135.

[13] C. von Wagner, G. Baio, R. Raine, J. Snowball, S. Morris, W. Atkin, A. Obichere, G. Handley, R.F. Logan, S. Rainbow, S. Smith, S. Halloran, J. Wardle, Inequalities in participation in an organized national colorectal cancer screening programme: results from the first 2.6 million invitations in England, Int. J. Epidemiol. 40 (2011) 712–8. doi:10.1093/ije/dyr008.

[14] S.H. Lo, S. Halloran, J. Snowball, H. Seaman, J. Wardle, C. von Wagner, Colorectal cancer screening uptake over three biennial invitation rounds in the English bowel cancer screening programme, Gut. 64 (2015) 282–291. doi:10.1136/gutjnl-2013-306144.

[15] A. Adler, S. Geiger, A. Keil, H. Bias, P. Schatz, T. DeVos, J. Dhein, M. Zimmermann, R. Tauber, B. Wiedenmann, Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany, BMC Gastroenterol. 14 (2014) 183. doi:10.1186/1471-230X-14-183.

[16] E.G. Liles, G.D. Coronado, N. Perrin, A.H. Harte, R. Nungesser, N. Quigley, N.T. Potter, G. Weiss, T. Koenig, T. DeVos, Uptake of a colorectal cancer screening blood test is higher than of a fecal test offered in clinic: a randomized trial, Cancer Treat. Res. Commun. 10 (2017) 27–31. doi:10.1016/J.CTARC.2016.12.004.

[17] J.E. Allison, R. Feldman, I.S. Tekawa, Hemoccult screening in detecting colorectal neoplasm: sensitivity, specificity, and predictive value. Long-term follow-up in a large group practice setting, Ann. Intern. Med. 112 (1990) 328–33.

[18] T.F. Imperiale, D.F. Ransohoff, S.H. Itzkowitz, B.A. Turnbull, M.E. Ross, Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population, N. Engl. J. Med. 351 (2004) 2704–14. doi:10.1056/NEJMoa033403.

[19] UK National Screening Committee, Note of the meeting held on the 19 November 2015 at Goodenough College, London, 2015.

[20] S. Moss, C. Mathews, T.J. Day, S. Smith, H.E. Seaman, J. Snowball, S.P. Halloran, Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England, Gut. 66 (2017) 1631–1644. doi:10.1136/gutjnl-2015-310691.

[21] U. Menon, A. Gentry-Maharaj, A. Ryan, A. Sharma, M. Burnell, R. Hallett, S. Lewis, A. Lopez, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, S. Skates, M. Parmar, I. Jacobs, Recruitment to multicentre trials—lessons from UKCTOCS: descriptive study, BMJ. 337 (2008) a2079–a2079. doi:10.1136/bmj.a2079.

[22] I.J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J.K. Kalsi, N.N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, D.N. Crump, S.K. Davies, A. Dawnay, S. Dobbs, G. Fletcher, J. Ford, K. Godfrey, R. Gunu, M. Habib, R. Hallett, J. Herod, H. Jenkins, C. Karpinskyj, S. Leeson, S.J. Lewis, W.R. Liston, A. Lopes, T. Mould, J. Murdoch, D. Oram, D.J. Rabideau, K. Reynolds, I. Scott, M.W. Seif, A. Sharma, N. Singh, J. Taylor, F. Warburton, M. Widschwendter, K. Williamson, R. Woolas, L. Fallowfield, A.J. McGuire, S. Campbell, M. Parmar, S.J. Skates, Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial, Lancet. 387 (2016) 945–956. doi:10.1016/S0140-6736(15)01224-6.

[23] S.J. Skates, D.K. Pauler, I.J. Jacobs, Screening Based on the Risk of Cancer Calculation From Bayesian Hierarchical Changepoint and Mixture Models of

Longitudinal Markers, J. Am. Stat. Assoc. 96 (2001) 429–439. doi:10.1198/016214501753168145.

[24] S.J. Skates, U. Menon, N. MacDonald, A.N. Rosenthal, D.H. Oram, R.C. Knapp, I.J. Jacobs, Calculation of the Risk of Ovarian Cancer From Serial CA-125 Values for Preclinical Detection in Postmenopausal Women, J. Clin. Oncol. 21 (2003) 206s–210. doi:10.1200/JCO.2003.02.955.

[25] H. Møller, S. Richards, N. Hanchett, S.P. Riaz, M. Lüchtenborg, L. Holmberg, D. Robinson, Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates, Br. J. Cancer. 105 (2011) 170–176. doi:10.1038/bjc.2011.168.

[26] A. Dregan, H. Møller, T. Murray-Thomas, M.C. Gulliford, Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study, Cancer Epidemiol. 36 (2012) 425–429. doi:10.1016/j.canep.2012.05.013.

Chapter 2. Background

The literature is clear: after adjustment for attendance (per protocol analysis), undertaking a one-off gFOBT reduces the risk of related mortality by 25% [1] and a sole flexible sigmoidoscopy by 43% [2]. CRC screening is predicated on detecting cancers before symptoms present—either during the premalignant phase where simple polypectomy is curative or at a localised stage with the expectation that these CRCs are highly treatable and have better outcomes. Screening tests are by no means diagnostic, however; and are instead intended to triage from an at-risk population the few with suspect signs of underlying disease who should undergo diagnostic colonoscopy.

Merely detection of asymptomatic cancers alone cannot justify population-based screening due to several biases associated with screening. That is the propensity for screening to detect indolent tumours likely to have better outcomes (a length-time bias), the detection of disease earlier with no prolonging of life (a lead-time bias) or the overdiagnosis and unnecessary medicalisation of cancers unlikely to cause harm within one's lifetime [3]. Recognising these challenges, Wilson and Junger published the *Principles and Practice of Screening for Disease* on behalf of the World Health Organisation in which ten principles fundamental for an efficacious screening programme were outlined [4]. These are listed verbatim:

—The condition sought should be an important health problem

—There should be an accepted treatment for patients with recognized disease, and treatment should be better at an earlier stage

—Facilities for diagnosis and treatment should be available

—There should be a recognizable latent or early symptomatic stage

—There should be a suitable test or examination

—The test should be acceptable to the population

—The natural history of the condition, including development from latent to declared disease, should be adequately understood

—There should be an agreed-upon policy on whom to treat as patients

—The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole

—Case-finding should be a continuing process and not a "once and for all" project.

The principles pertinent to the current work will herein be discussed in the context of the current literature.

2.1 Morbidity and Mortality

'The condition sought should be an important health problem.'

World Health Organisation [4]

CRC is a considerable burden—the fourth commonest and second-most fatal of all cancers in the UK. During 2015, British cancer registries recorded 41,804 morbidities and 16,384 related-mortalities [5]. An estimated 76% of CRC patients survive one-year and 59 & 57% are expected to survive through five & ten years, respectively [6].

Incidence is a function of demographics—size, age structure—and patterns in the exposure to risk or protective factors. Thus, age-standardised rates are useful for monitoring progress in healthcare interventions in real terms. The age-standardised incidence rate (per 100,000 persons) within the UK has remained more-or-less stable (–0.6% (69.9/70.3)) since 1993, but when stratified by gender amounts to a 1.0% (57.2/57.8) & 3.3% (85.3/88.2) decline within females and males during the same period (Figure 2–1) [5]. Age-standardised incidence rates are projected to be 6.6% lower in women and 15.1% lower in men come 2035 [7]. An estimated 58,176 CRCs will be diagnosed during 2030 [7].

Mortality is a function of incidence and the success of public healthcare. The age-standardised mortality rate (per 100,000) for the UK has decreased by 33.8% over 1993–2015; and by 35.5% (21.1/32.7) for women and by 34.3% (33.1/50.4) for men during the same period (Figure 2–1) [8]. A fundamental driver underlying this decrease is the that a higher proportion are surviving for longer: the proportion of English and Welsh cancer patients surviving at least five years in 2010–11, for example, has increased by 34% since 1971–1972 and by 5% since 2005–06 [6]. The age-standardised mortality rates are projected to continue decreasing and come 2035 will be 18.8% and 28.6% lower for women & men, respectively [7].

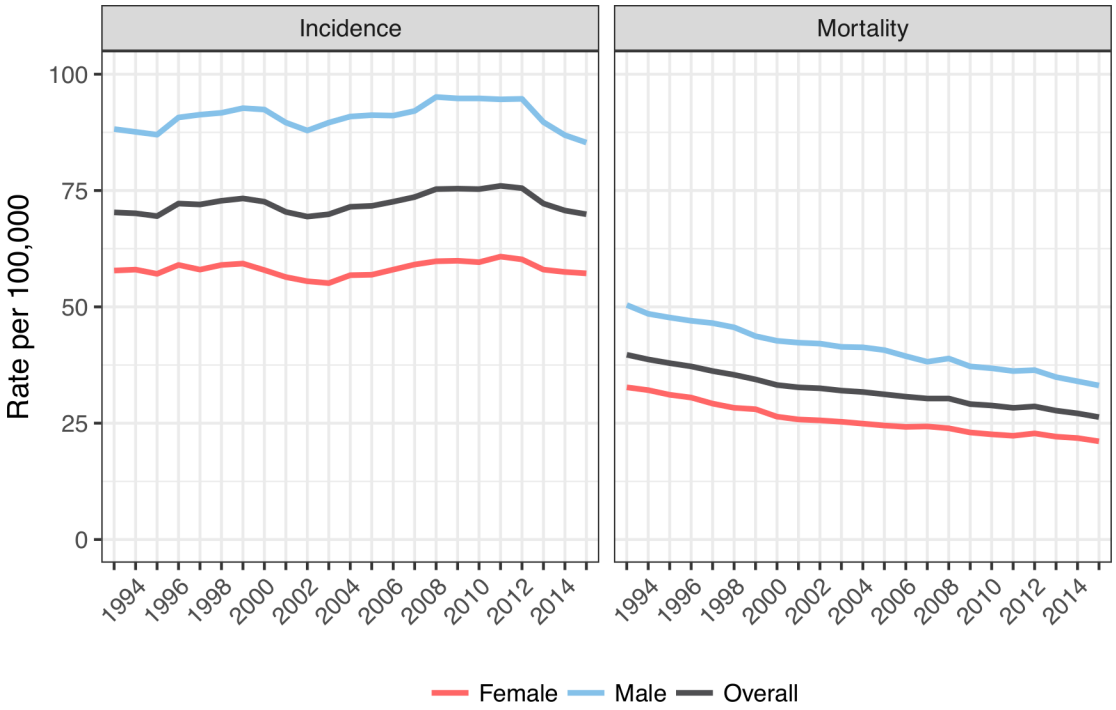


Figure 2–1: European age-standardised incidence and mortality rates in the UK during 1993–2015 [5,8]. Overall incidence and mortality rates decreased by 0.6% (69.9/70.3) and 33.8% (26.3/39.7) during 1993–2015, respectively.

It is immediately clear that there exists disparities in the incidence of and mortality from CRC according to gender. This also holds true for socioeconomic development nationally and internationally. Age-standardised incidence rates within transitioned (females 20.9, males 30.3 per 100,000) are consistently higher than within transitioning countries (females 5.9, males 8.4 per 100,000) [9], while migrants' risk rapidly acclimatises toward that of natives [10]. Within the UK, an analysis of 21,114 CRCs recorded by Scottish cancer registries found an emergence of a socioeconomic gradient in women and the widening of an existing gradient in men during 2001–12 [11]. And within England, incidence correlates with income for men but not women [12].

In a landmark study on the heritability of cancer in identical and non-identical twins, the non-genetic component was attributed with 65% of the variance in CRC susceptibility [13]. Thus, it is acknowledged that it is differences in the prevalence of exposure to environmental risk factors that accounts for the variability in susceptibility between females & males [14] and affluence & poverty [15–17]. Notable among these risk factors are diets too rich in red meat, sedentary lifestyles likely leading to overweightness and obesity, and use of tobacco and alcohol, which altogether were attributed with 54% of all CRCs diagnosed in the UK during 2015 [18]. It is likely that, given that the world's poorest have some of the lowest incidence rates worldwide, poverty in transitioning

countries is initially a barrier to exposure to risk factors, but within transitioned countries is a means of escaping health inequalities. Of the 35% of risk susceptibility attributed to heritable components, an estimated 3–5% of diagnoses are predisposed by the inheritance of highly penetrant susceptibility genes [19], and the remaining variance is likely to be uncovered by genome-wide association studies, which have hitherto identified 45 low-susceptibility gene variants associated with a slight (ORs 1.05–1.53) but frequent (7–91% prevalence in Europeans) increase to risk [20,21].

2.2 Natural History

'There should be a recognisable latent or early symptomatic stage. The natural history of the condition, including development from latent to declared disease, should be adequately understood.'

World Health Organisation [4]

Cancer is a disease of inappropriate cell proliferation—a pathological mitosis. A cancerous cell is a lineal descendant of a once-normal cell that, over many generations, has accumulated a succession of slight variations by mutability in the genome; each mutant variation naturally selected by proving profitable to a cancer cell's survival. The divergence of cells from the highly regulated pathways governing cell division is often preceded by mutations that render oncogenes a dominant gain-of-function or tumour suppressor genes with a recessive loss of function [22,23]. More specifically, survival descends from commandeering normal physiological processes—rationalised as the Hallmarks of Cancer: sustained growth signalling, evasion of growth suppressors, resistance to apoptosis, sustenance through angiogenesis, bypassing senescence, and mobility through invasion and metastasis [24,25].

Despite the remarkable diversity in the mutability of CRCs, they invariably progress through a finite number of molecular pathways involving a sequential accumulation of common genes altered in a high proportion of tumours [26–28]. Nonetheless, Bert Vogelstein observed in 1988 that the 'molecular alterations accumulated in a fashion that paralleled the clinical progression of tumours' [29]. Termed in namesake the Vogelstein model, it is the paradigm favoured by most on how normal colorectal epithelium transitions to benign adenoma and eventually to adenocarcinoma. Few benign adenomas will to transition to cancer, however, and there is robust evidence for advanced adenomas—those larger than 1 cm with high-grade dysplasia and villous

morphology—being the most clinically important precursor to CRC [30,31]. An estimated 2.6% and 4.2% benign adenomas progress to carcinomas in one's fifties and late sixties [32]. Nonetheless, excluding indolent and fast-growing cancers, the phase of premalignancy between a large adenoma to adenocarcinoma is thought to be lengthy at 10–17 years [30,33]. This provides ample opportunity for early detection and intervention.

Breach of the basement membrane demarcates the clinical progression from benign adenoma to CRC. Most diagnoses, barring those markedly advanced and subject to palliative care, will undergo clinical staging to aid in managing healthcare. Staging takes form of a histological assessment of the cancerous spread outwards from the luminal epithelial and involvement of resected lymph nodes combined with imaging technologies to detect any metastatic spread. This thesis will use the modified four-stage system first proposed by Cuthbert Duke wherein growth up to but not through the muscularis propria demarcates Dukes A, penetration of the serosa and peritonea as Dukes B, invasion of regional lymphatic nodes as Dukes C, and metastasis to distant organs as Dukes D [34,35]. (Use of the TNM or AJCC systems in the literature are converted to the Dukes equivalent). During 1996–2002, before implementation of population-based screening in England, 93, 77, 48, and 7% of patients diagnosed with Dukes A–D CRCs inclusively survived five years [36]. And yet, excluding the third of cases for whom a clinical stage was not determined, 13, 37, 36, and 14% of CRCs were diagnosed at Dukes A–D inclusively [36], thereby suggesting that symptomatic presentation of CRC occurs far too late to be relied on for many patients.

2.3 Faecal Occult Blood

'There should be a suitable test or examination.
The test should be acceptable to the population.'
World Health Organisation [4]

Since the dawn of the Twentieth century, it was known that both benzidine and guaiac—a phenolic resin—could detect microscopic traces of blood [37]. It was not until 1967 though that, suspecting CRCs leaked occult blood detectable on stools, David Greegor, a medical internist from Ohio, would fashion together what is now recognisable as the gFOBT [38]. In the presence of haemoglobin—and more specifically, the haem moiety—and hydrogen peroxide, guaiac is oxidised into a detectable blue-coloured compound indicative of a positive finding [39]. Testing on 2,000 patients undergoing routine

examination, each showing no symptoms suggestable of a diagnosis, Greegor found that seven under his care tested positive for occult blood. All seven, it later turned out, had CRC. Smith Kline Diagnostics soon after marketed the test as Haemoccult.

In the wake of considerable interest in the use of the gFOBT for CRC screening, numerous randomised-controlled trials commenced, all aiming to discern a mortality benefit—the utmost important endpoint that more-or-less rules out screening biases. The first, the Minnesota Colon Cancer Control study, during 1976–77 randomised 46,551 men and women aged 40–80 to a control arm or were offered gFOBT-based screening with diagnostic colonoscopy or double-contrast barium enema follow-up for positive results [40]. The primary endpoint was mortality from CRC. Trials began soon after in the UK [41], Sweden [42], and then Denmark [43], screening subjects aged 45–74, 60–64, and 45–75-years, respectively, each intent on comparing related mortality within screenees and controls. Trial results were published during 1999–2008—all bar the Funen trial (Denmark) unanimous in reporting a 13–21% reduction in related mortality due to biennial screening (Table 2–1). Despite the equivocal 11% (95% CI (0.78–1.01) reduction reported for biennial gFOBT in Funen, the manuscript reported a 43% reduced mortality in those who attended all nine screening rounds and a 16% reduction for those who received the gFOBT annually [44]. In a 2008 meta-analysis of these four trials (n 327,043), biennial screening with the gFOBT was altogether attributed with a 15% mortality reduction (RR 0.85 (95% CI 0.78–0.92)), even despite unfulfilled uptake [1]. The reduction in mortality after adjustment for attendance to at least one screening episode was 25% (RR 0.75 (95% CI 0.66–0.84)) [1]. Moreover, the secondary analysis found a 5–17% increase in the proportion of highly treatable Dukes A CRCs within screenees relative to controls [1].

Table 2–1: An overview of randomised controlled trials studying the impact of gFOBT-based screening on CRC mortality.

Study	Screening rounds	Years follow-up	Mortality reduction RR (95% CI)
Biennial gFOBT			
Funen, Denmark [44]	9	17.0 (8.5–18.4)	0.89 (0.78–1.01)
Goteborg, Sweden [45]	2	15.8 (11.3–19.5)	0.84 (0.71–0.99)
Minnesota, USA [46]	6	18 (NA–NA)	0.79 (0.62–0.97)
Nottingham, UK [47]	6	11.7 (8.4–18.4)	0.87 (0.78–0.97)
Annual gFOBT			
Funen, Denmark [44]	9	17.0 (8.5–18.4)	0.84 (0.73–0.96)
Minnesota, USA [46]	11	18 (NA–NA)	0.67 (0.51–0.83)

Crucial to any efficacious screening programme is a viable screening test. A sought-after screening test would correctly identify all those destined for future cancer diagnosis while ruling out all those undestined from unnecessary diagnostic workup. Rarely—if ever—is this the reality, however. Throughout the literature, test performances are communicated in terms of the proportion of screenees with an underlying disease that return a positive test (i.e. the sensitivity) and the proportion of screenees with no underlying disease who return a negative test (i.e. the specificity). Thus, a test with high sensitivity has a low likelihood of returning false-negative findings; and a test with high specificity has a low likelihood of returning false-positive findings.

Consider as an example a scenario where 1,000 apparently well persons showing no symptoms suggestable of disease were screened with a test having a sensitivity of 50% and a specificity of 98%. Irregardless of the screening findings, all underwent colonoscopy; revealing 900 without CRC and 100 with CRC. The test therefore returns 50 true positive findings for whom can undergo immediate treatment and 882 true negative findings for whom there are no causes for concern and are thereafter invited to undergo screening in the next cycle (Table 2–2). On the other hand, the test also returns 50 false negative for whom may receive delayed treatment and 50 false negative findings for whom undergo unnecessary diagnostic colonoscopy at cost to personal health and burden to public health endoscopic resources (Table 2–2).

Table 2–2: Confusion matrix classifying findings on the colonoscopy outcomes and screening positivity.

	Colonoscopy +	**Colonoscopy –**
Screen +	50	25
Screen –	50	102

It is clear that screening programmes favour tests having both high sensitivity and specificity. With few exceptions, however, increasing the technical sensitivity by lowering the threshold demarcating test positivity is at a trade-off to specificity—that a higher proportion of well screenees are falsely positive. Vice versa, increasing the technical specificity by raising the threshold demarcating test positivity is at a trade-off to sensitivity—that a higher proportion of screenees with underlying disease are falsely negative. This trade-off between technical sensitivity and specificity is often communicated as the Area Under the Curve (AUC), which describes the probability that a screening score is higher for a randomly sampled screenee with underlying disease than for a random screenee without disease. A population-based screening programme where

disease is a rarity requires a test with high specificity to minimise the harmful screens and avoid overwhelming diagnostic work-up services that result from false-positive findings.

Mortality benefit withstanding, the gFOBT is far from perfect. The gFOBT is able to detect only a proportion of asymptotic CRCs (Table 2–3) and right-sided cancers prevalent in women [48–52], which could provide false reassurance and delayed treatment for patients with these cancers. (The proportion of right-sided CRCs in the UK for women & men during 2010–12 was 72 & 61%, respectively [5] and 73% for all CRCs diagnosed in England and Wales during 1971–1994 [53]). In screenees with no CRC in retrospect, false positivity may be triggered due to bleeding associated with haematological conditions and other maladies of the gastrointestinal track [54]. It has also thought that the psuedoperoxidase reaction could be triggered animal blood contained in red meats or by the peroxidases contained in certain vegetables, although a meta-analysis found no association between dietary restrictions and positivity rates [55].

Table 2–3: Sensitivity of the gFOBT for detecting asymptomatic CRCs as a function of lead time [56].

Lead time	Benign adenoma (95% CIs)	CRC (95% CIs)
≤ 1 year	0.36 (0.27–0.45)	0.50 (0.33–0.67)
≤ 2 years	0.28 (0.20–0.35)	0.43 (0.29–0.58)
≤ 4 years	0.17 (0.13–0.22)	0.25 (0.16–0.34)

2.4 The NHS Bowel Cancer Screening Programme

On the successful implementation of biennial gFOBT-based screening in pilot sites nested within routine NHS healthcare in England and Scotland and replication of findings from the Nottingham trial [57–59], the BCSP was formally recommended by the UK National Screening Committee and began offering organised, biennial population-based screening with the gFOBT for the 60–69-years-old demographic. This age range chosen as a pragmatic balance between the age-specific prevalence, potential years of life saved, and avoidance of unnecessary treatment of overdiagnosed CRCs in the elderly (Figure 2–2) [60]. (This has since extended to 50–74 in 2018 [61], though Scotland has offered screening for 50–74-years for some time). The gFOBT kit used consists of six test windows impregnated with guaiac, which are completed at home with three separate bowel movements and mailed to the laboratory for processing. (The BCSP does not require any dietary restrictions). Generally, 2% of screenees will have detectable occult blood—defined as 5–6 positive windows on their first test or further positivity in any of

the subsequent two retake tests following inconclusive findings (1–4 positive windows)—are thereafter offered diagnostic colonoscopy [62]. During the initial prevalence round of the BCSP, an approximate 627 and 609 gFOBTs were completed for each advanced adenoma and CRC detected, respectively [62].

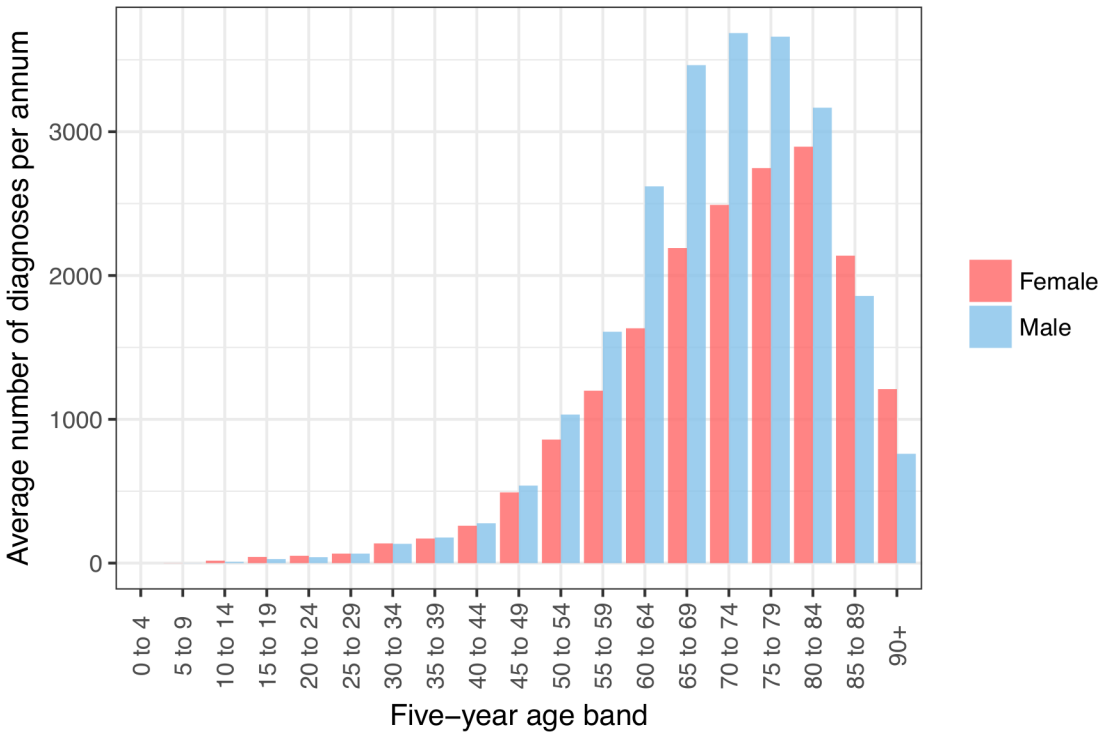


Figure 2–2: Age distribution of CRC incidences in the UK during 2013–15 [5].

Since implementation of the BCSP, a tenth of all CRCs are now diagnosed via screening (one-year survival rate (98% (95% CI 97–98)), and the proportion of CRCs diagnosed after emergency admission (one-year survival rate 50% (95% CI 49–51)) has decreased by 4 percentage points during 2006–16 [63,64]. Moreover, in alignment with the findings of clinical trials, a greater proportion of screen-detected CRCs are diagnosed at Dukes A and fewer at Dukes D relative to those detected perchance [49–51].

Nonetheless, the potential benefits of the BCSP is undermined by low uptake—particularly relative to the longstanding NHS screening programmes for breast (71%) [65] and cervical cancers (71%) [66]. In England, for example, just 54% of those eligible at initial rollout during 2006–08 returned a completed test kit [62] and the proportion of first-time invitees engaging in the BCSP has declined year on year from 53% during 2010 to 49% in 2015 [67]. (Uptake of first-time and repeated invitees in Scotland was 56% during 2007–14 [68]). Inequalities in uptake have existed since implementation of the BCSP within areas of socioeconomic deprivation (35% in the most vs 61% in the least-

deprived quintiles) and ethnic diversity (38% in the most vs. 55% in the least diverse areas), while barriers to uptake also appear gender-specific (51% of males vs. 56% of females) [69]. These patterns of inequalities in uptake are still apparent at the time of writing [67].

The BCSP faces a number of challenges unencountered by the current NHS screening programmes for breast and cervical cancers. Indeed, focus groups of non-acceptors commonly cited their objection to handling stools and their perceived detachment from a formal clinical setting as reasons for their non-acceptance [70]. (By comparison, the mammography and cytology/HPV tests used for breast and cervical screening, respectively, are undertaken by a professional in a formal clinical setting). Busyness, too, may cause delayed completion and inadvertent refusal [70]. Complete uptake, however, is rarely realistic and invitees' autonomy in making an informed choice on whether to undertake screening must be respected [70–72].

Prior engagement in screening is a strong predictor of future participation [73]. And continued invitation of previous non-responders may also prompt subsequent participation [74]. A number of interventions designed to equalise the socioeconomic-gradient to screening uptake have since been trailed with varying degrees of success. Those strategies that proved successful were a reminder letter sent to non-responding invitees to the BCSP (with notable effect in promoting uptake within the most-deprived) [75,76], a letter of endorsement from their registered GP [75,77], an advisory telephone call at time of invitation [78], and—of interest to the present work—provision of the next-generation FIT [79,80], which in 2015 the UK National Screening Committee recommended it replace the gFOBT [81].

Like the its predecessing gFOBT, the FIT detects trace levels of blood in stools but in contrast relies on its detection via a labelled antibody targeting the globin component of *human* haemoglobin. Thus, the FIT has numerous advantageous over the gFOBT; not least being unaffected by diet and medication, being able to detect blood at lower concentrations (10 μg haemoglobin g^{-1} faeces for FIT vs. 600 μg haemoglobin g^{-1} faeces for gFOBT [as stated in manufacturers' literature], and requiring just one stool sample as opposed the three required for the gFOBT. These have translated to improved outcomes within the context of existing screening programmes. Provision of a pilot FIT within the BCSP in England increased uptake by 7 percentage points (64% for FIT vs. 59% for gFOBT), and most notably bettered uptake within previous non-responders and men [80]. And in Scotland, acceptance for the FIT was 5 percentage points higher (59%)

than for the gFOBT (54%) alongside a dramatic increase in involvement of men and those of high social deprivation [79]. Moreover, a considerable proportion of gFOBT false positives later submitted for the FIT returned negative, thereby implying its use can ease the demand for unnecessary colonoscopies [82].

The quantitative readout in terms of μg haemoglobin g^{-1} faeces also provides ample opportunity for risk-tailored screening, which was difficult to implement with the qualitative readout based on window positivity and repeat testing the gFOBT provides. Stakeholders are thus able to adjust the threshold demarcating test positivity to match existing endoscopic services. It is also reported that the stool haemoglobin concentration correlates with future risk of diagnosis and the severity of disease [83,84]. Accordingly, the stool haemoglobin content of past screening episodes may be useful in determining the need, frequency, and modality of subsequent screens [85,86].

2.5 The NHS Bowel Scope Screening Programme

Owing to their indolent natural history, most left-sided CRCs that go on to cause harm have developed to at least a benign adenoma by the age of 60 [87]. Moreover, their removal substantially lowers the long-term risk of developing CRC [88] and, unlike diagnostic colonoscopy, flexible sigmoidoscopy takes considerably less time and can be undertaken by a nurse endoscopist [89].

On these observations, a randomised controlled trial was commissioned in the UK to evaluate the impact a one-off endoscopic screen had on CRC incidence and related-mortality [90]. During 1994–99, 170,432 men and women aged 55–64 years were recruited from NHS trusts in England, Scotland and Wales. Of these, 170,432 were to receive their standard care with no screening as baseline controls, while the remaining 57,237 were to receive a once-only flexible sigmoidoscopy screen with immediate polypectomy of small polyps and further colonoscopy for subjects with high-risk polys or suspected CRC. After an average 11-years follow-up, the trial reported that once-only flexible sigmoidoscopy on an intention-to-treat analysis reduced incidence of CRC by 23% (HR 0.77 (95% CI 0.70–0.84)) and mortality by 31% (HR 0.69 (95% CI 0.59–0.82)) [2]. After adjustment for the 71% whom attended screening (per-protocol analysis), CRC incidence and mortality was reduced by 33% (HR 0.67 (95% CI 0.60–0.76)) and 43% (HR 0.57 (95% CI 0.45–0.72)), respectively. No gender-specific difference in the morbidity benefit were reported despite the lower prevalence of left-sided CRCs in women. Elsewhere, and based on the same protocol, the SCORE trial undertaken in Italy reported

a reduction in morbidity (HR 0.69 (95% 0.56–0.86)) and mortality (HR 0.62 (95% CI 0.40–0.86)) per protocol after an average of 11 years [91].

Informed by these results, the UK National Screening Committee recommended the provision of a one-off flexible sigmoidoscopy screen [92] and this was commissioned for English residents at age 55 years within the NHS Bowel Scope Programme. Nevertheless, endoscopic capacity has hindered its implementation and its long-term future is dependent on its proven benefits in combination with the revised provision of the FIT to the 50–74-year old demographic [61].

2.6 References

[1] P. Hewitson, P. Glasziou, E. Watson, B. Towler, L. Irwig, Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update, Am J Gastroenterol. 103 (2008) 1541–1549. doi:10.1111/j.1572-0241.2008.01875.x.

[2] W.S. Atkin, R. Edwards, I. Kralj-Hans, K. Wooldrage, A.R. Hart, J.M. Northover, D.M. Parkin, J. Wardle, S.W. Duffy, J. Cuzick, Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial, Lancet. 375 (2010) 1624–1633. doi:10.1016/S0140-6736(10)60551-X.

[3] J.M. Croswell, D.F. Ransohoff, B.S. Kramer, Principles of cancer screening: lessons from history and study design issues, Semin. Oncol. 37 (2010) 202–15. doi:10.1053/j.seminoncol.2010.05.006.

[4] J.M.G. Wilson, G. Jungner, Principles and Practice of Screening for Disease, World Heal. Organ. Public Heal. Pap. 34 (1968).

[5] Cancer Research UK, Bowel cancer incidence statistics [UK, 2015, ICD-10 C18–C20], (2018). http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence (accessed February 28, 2018).

[6] Cancer Research UK, Bowel Cancer (C18-C20 and C21.8), Age-Standardised One-, Five- and Ten-Year Net Survival, Adults (Aged 15-99), England and Wales, 2010-2011, (2014). https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival (accessed November 17, 2018).

[7] C.R. Smittenaar, K.A. Petersen, K. Stewart, N. Moitt, Cancer incidence and mortality projections in the UK until 2035, Br. J. Cancer. 115 (2016) 1147–1155. doi:10.1038/bjc.2016.304.

[8] Cancer Research UK, Bowel cancer mortality statistics [UK, 2015, ICD-10 C18–C20], (2019).

[9] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA. Cancer J. Clin. 68 (2018) 394–424. doi:10.3322/caac.21492.

[10] S.M. Mousavi, M. Fallah, K. Sundquist, K. Hemminki, Age- and time-dependent changes in cancer incidence among immigrants to Sweden: colorectal, lung, breast and prostate cancers, Int. J. Cancer. 131 (2012) E122–E128. doi:10.1002/ijc.27334.

[11] E.J. Tweed, G.M. Allardice, P. McLoone, D.S. Morrison, Socio-economic inequalities in the incidence of four common cancers: a population-based registry study, Public Health. 154 (2018) 1–10. doi:10.1016/j.puhe.2017.10.005.

[12] National Cancer Intelligence Network, Cancer by Deprivation in England, 2014.

[13] P. Lichtenstein, N. V Holm, P.K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, K. Hemminki, Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland, N Engl J Med. 343 (2000) 78–85. doi:10.1056/NEJM200007133430201.

[14] D.M. Parkin, A.H. Olsen, P. Sasieni, The potential for prevention of colorectal cancer in the UK, Eur J Cancer Prev. 18 (2009) 179–190. doi:10.1097/CEJ.0b013e32830c8d83.

[15] R. Hiscock, L. Bauld, A. Amos, S. Platt, Smoking and socioeconomic status in England: the rise of the never smoker and the disadvantaged smoker, J. Public Health (Oxf). 34 (2012) 390–6. doi:10.1093/pubmed/fds012.

[16] S.V. Katikireddi, E. Whitley, J. Lewsey, L. Gray, A.H. Leyland, Socioeconomic status as an effect modifier of alcohol consumption and harm: analysis of linked cohort data, Lancet Public Heal. 2 (2017) e267–e276. doi:10.1016/S2468-2667(17)30078-6.

[17] D. Bann, W. Johnson, L. Li, D. Kuh, R. Hardy, Socioeconomic Inequalities in Body Mass Index across Adulthood: Coordinated Analyses of Individual Participant Data from Three British Birth Cohort Studies Initiated in 1946, 1958 and 1970, PLoS Med. 14 (2017) e1002214. doi:10.1371/journal.pmed.1002214.

[18] K.F. Brown, H. Rumgay, C. Dunlop, M. Ryan, F. Quartly, A. Cox, A. Deas, L. Elliss-Brookes, A. Gavin, L. Hounsome, D. Huws, N. Ormiston-Smith, J. Shelton, C. White, D.M. Parkin, The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015, Br. J. Cancer. 118 (2018) 1130–1141. doi:10.1038/s41416-018-0029-6.

[19] S.R. Cairns, J.H. Scholefield, R.J. Steele, M.G. Dunlop, H.J.W. Thomas, G.D. Evans, J.A. Eaden, M.D. Rutter, W.P. Atkin, B.P. Saunders, A. Lucassen, P. Jenkins, P.D. Fairclough, C.R.J. Woodhouse, Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002), Gut. 59 (2010) 666–89. doi:10.1136/gut.2009.179804.

[20] M.A. Jenkins, E. Makalic, J.G. Dowty, D.F. Schmidt, G.S. Dite, R.J. MacInnis, D. Ait Ouakrim, M. Clendenning, L.B. Flander, O.K. Stanesby, J.L. Hopper, A.K. Win, D.D. Buchanan, Quantifying the utility of single nucleotide polymorphisms to guide colorectal cancer screening, Future Oncol. 12 (2016) 503–13. doi:10.2217/fon.15.303.

[21] R.S. Houlston, COGENT (COlorectal cancer GENeTics) revisited, Mutagenesis. 27 (2012) 143–51. doi:10.1093/mutage/ger059.

[22] D. Stehelin, H.E. Varmus, J.M. Bishop, P.K. Vogt, DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA, Nature. 260 (1976) 170–173. doi:10.1038/260170a0.

[23] A.G. Knudson, Mutation and cancer: statistical study of retinoblastoma, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 820–3.

[24] D. Hanahan, R.A. Weinberg, The hallmarks of cancer, Cell. 100 (2000) 57–70.

[25] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell. 144 (2011) 646–74. doi:10.1016/j.cell.2011.02.013.

[26] E.R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis, Cell. 61 (1990) 759–67.

[27] L.D. Wood, D.W. Parsons, S. Jones, J. Lin, T. Sjöblom, R.J. Leary, D. Shen, S.M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezso, V. Ustyanksky, T. Nikolskaya, Y. Nikolsky, R. Karchin, P.A. Wilson, J.S. Kaminker, Z. Zhang, R. Croshaw, J. Willis, D. Dawson, M. Shipitsin, J.K. V Willson, S. Sukumar, K. Polyak, B.H. Park, C.L. Pethiyagoda, P.V.K. Pant, D.G. Ballinger, A.B. Sparks, J. Hartigan, D.R. Smith, E. Suh, N. Papadopoulos, P. Buckhaults, S.D. Markowitz, G. Parmigiani, K.W. Kinzler, V.E. Velculescu, B. Vogelstein, The genomic landscapes of human breast and colorectal cancers, Science. 318 (2007) 1108–13. doi:10.1126/science.1145720.

[28] B. Vogelstein, N. Papadopoulos, V.E. Velculescu, S. Zhou, L.A. Diaz, K.W. Kinzler, Cancer Genome Landscapes, Science (80-.). 339 (2013) 1546–1558. doi:10.1126/science.1235122.

[29] B. Vogelstein, E.R. Fearon, S.R. Hamilton, S.E. Kern, A.C. Preisinger, M. Leppert, Y. Nakamura, R. White, A.M. Smits, J.L. Bos, Genetic alterations during colorectal-tumor development, N. Engl. J. Med. 319 (1988) 525–32. doi:10.1056/NEJM198809013190901.

[30] T. Muto, H.J. Bussey, B.C. Morson, The evolution of cancer of the colon and rectum, Cancer. 36 (1975) 2251–70.

[31] H. Shinya, W.I. Wolff, Morphology, anatomic distribution and cancer potential of colonic polyps, Ann. Surg. 190 (1979) 679–83.

[32] H. Brenner, M. Hoffmeister, C. Stegmaier, G. Brenner, L. Altenhofen, U. Haug, Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies, Gut. 56 (2007) 1585–9. doi:10.1136/gut.2007.122739.

[33] S. Jones, W.-D. Chen, G. Parmigiani, F. Diehl, N. Beerenwinkel, T. Antal, A. Traulsen, M.A. Nowak, C. Siegel, V.E. Velculescu, K.W. Kinzler, B. Vogelstein, J. Willis, S.D. Markowitz, Comparative lesion sequencing provides insights into tumor evolution, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 4283–8. doi:10.1073/pnas.0712345105.

[34] C.E. Dukes, The classification of cancer of the rectum, J. Pathol. Bacteriol. 35 (1932) 323–332. doi:10.1002/path.1700350303.

[35] National Collaborating Centre for Cancer (UK), Colorectal Cancer: The Diagnosis and Management of Colorectal Cancer, National Collaborating Centre for Cancer (UK), 2011.

[36] National Cancer Intelligence Network, Colorectal Cancer Survival by Stage, (2009). http://www.ncin.org.uk/publications/data_briefings/colorectal_cancer_survival_by_stage (accessed October 24, 2013).

[37] G.H.M. Thornton, D.G. Illingworth, An Evaluation of the Benzidine Test for Occult Blood in the Feces, Gastroenterology. 28 (1955) 593–605. doi:10.1016/S0016-5085(55)80045-5.

[38] D.H. Greegor, Diagnosis of large-bowel cancer in the asymptomatic patient, JAMA. 201 (1967) 943–5.

[39] J.F. Kratochvil, R.H. Burris, M.K. Seikel, J.M. Harkin, Isolation and characterization of α-guaiaconic acid and the nature of guaiacum blue, Phytochemistry. 10 (1971) 2529–2531. doi:10.1016/S0031-9422(00)89901-X.

[40] V.A. Gilbertsen, T.R. Church, F.J. Grewe, J.S. Mandel, R.B. McHugh, L.M. Schuman, S.E. Williams, The design of a study to assess occult-blood screening for colon cancer, J. Chronic Dis. 33 (1980) 107–114. doi:https://doi.org/10.1016/0021-9681(80)90034-X.

[41] J.D. Hardcastle, J.O. Chamberlain, M.H. Robinson, S.M. Moss, S.S. Amar, T.W. Balfour, P.D. James, C.M. Mangham, Randomised controlled trial of faecal-occult-blood screening for colorectal cancer, Lancet. 348 (1996) 1472–1477. doi:10.1016/s0140-6736(96)03386-7.

[42] J. Kewenter, S. Björk, E. Haglind, L. Smith, J. Svanvik, C. Aåhrén, Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27,700 subjects, Cancer. 62 (1988) 645–651. doi:10.1002/1097-0142(19880801)62:3<645::AID-CNCR2820620333>3.0.CO;2-#.

[43] O. Kronborg, C. Fenger, O. Søsndergaard, K.M. Pedersen, J. Olsen, Initial Mass Screening for Colorectal Cancer with Fecal Occult Blood Test: A Prospective Randomized Study at Funen in Denmark, Scand. J. Gastroenterol. 22 (1987) 677–686. doi:10.3109/00365528709011142.

[44] O. Kronborg, O.D. Jørgensen, C. Fenger, M. Rasmussen, Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds, Scand. J. Gastroenterol. 39 (2004) 846–51. doi:10.1080/00365520410003182.

[45] E. Lindholm, H. Brevinge, E. Haglind, Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer, Br. J. Surg. 95 (2008) 1029–36. doi:10.1002/bjs.6136.

[46] J.S. Mandel, T.R. Church, F. Ederer, J.H. Bond, Colorectal Cancer Mortality: Effectiveness of Biennial Screening for Fecal Occult Blood, JNCI J. Natl. Cancer Inst. 91 (1999) 434–437. doi:10.1093/jnci/91.5.434.

[47] J.H. Scholefield, S. Moss, F. Sufi, C.M. Mangham, J.D. Hardcastle, Effect of faecal occult blood screening on mortality from colorectal cancer: results from a randomised controlled trial, Gut. 50 (2002) 840–4.

[48] H. Brenner, U. Haug, S. Hundt, Sex Differences in Performance of Fecal Occult Blood Testing, Am. J. Gastroenterol. 105 (2010) 2457–2464. doi:10.1038/ajg.2010.301.

[49] R.J.C. Steele, P. McClements, C. Watling, G. Libby, D. Weller, D.H. Brewster, R. Black, F.A. Carey, C.G. Fraser, Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site, Gut. 61 (2012) 576–581. doi:10.1136/gutjnl-2011-300535.

[50] M.D. Gill, M.G. Bramble, C.J. Rees, T.J.W. Lee, D.M. Bradburn, S.J. Mills, Comparison of screen-detected and interval colorectal cancers in the Bowel Cancer Screening Programme, Br. J. Cancer. 107 (2012) 417–21. doi:10.1038/bjc.2012.305.

[51] E.J.A. Morris, L.E. Whitehouse, T. Farrell, C. Nickerson, J.D. Thomas, P. Quirke, M.D. Rutter, C. Rees, P.J. Finan, J.R. Wilkinson, J. Patnick, A retrospective observational study examining the characteristics and outcomes of tumours diagnosed within and without of the English NHS Bowel Cancer Screening Programme, Br. J. Cancer. 107 (2012) 757–764. doi:10.1038/bjc.2012.331.

[52] S.M. Moss, C. Campbell, J. Melia, D. Coleman, S. Smith, R. Parker, P. Ramsell, J. Patnick, D.P. Weller, Performance measures in three rounds of the English bowel cancer screening pilot, Gut. 61 (2011) 101–107. doi:10.1136/gut.2010.236430.

[53] D. Hayne, R.S. Brown, M. McCormack, M.J. Quinn, H.A. Payne, P. Babb, Current trends in colorectal cancer: site, incidence, mortality and survival in England and Wales, Clin. Oncol. (R. Coll. Radiol). 13 (2001) 448–52.

[54] E. He, R. Alison, R. Blanks, K. Pirie, G. Reeves, R.L. Ward, R. Steele, J. Patnick, K. Canfell, V. Beral, J. Green, Association of ten gastrointestinal and other medical conditions with positivity to faecal occult blood testing in routine screening: a large prospective study of women in England, Int. J. Epidemiol. (2019). doi:10.1093/ije/dyy271.

[55] M. Pignone, M.K. Campbell, C. Carr, C. Phillips, Meta-analysis of dietary restriction during fecal occult blood testing, Eff. Clin. Pract. 4 (n.d.) 150–6.

[56] J.E. Allison, R. Feldman, I.S. Tekawa, Hemoccult screening in detecting colorectal neoplasm: sensitivity, specificity, and predictive value. Long-term follow-up in a large group practice setting, Ann. Intern. Med. 112 (1990) 328–33.

[57] The UK CRC Screening Pilot Evaluation Team, Evaluation of the UK Colorectal Cancer Screening Pilot, 2003.

[58] R.J. Steele, R. Parker, J. Patnick, J. Warner, C. Fraser, N.A. Mowat, J. Wilson, F.E. Alexander, J.G. Paterson, A demonstration pilot trial for colorectal cancer screening in the United Kingdom: a new concept in the introduction of healthcare strategies, J. Med. Screen. 8 (2001) 197–202.

[59] UK Colorectal Cancer Screening Pilot Group, Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom, BMJ. 329 (2004) 133. doi:10.1136/bmj.38153.491887.7C.

[60] M.R. Law, J.K. Morris, N.J. Wald, The importance of age in screening for cancer, J. Med. Screen. 6 (1999) 16–20. doi:10.1136/jms.6.1.16.

[61] UK National Screening Committee, Note of the meeting held on the 29 June 2018 in Edinburgh, Scotland, 2018.

[62] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut. 61 (2012) 1439–46. doi:10.1136/gutjnl-2011-300843.

[63] Public Health England, Routes to Diagnosis 2006-2016 year breakdown version 2.0, 2018.

[64] L. Elliss-Brookes, S. McPhail, A. Ives, M. Greenslade, J. Shelton, S. Hiom, M. Richards, Routes to diagnosis for cancer – determining the patient journey using multiple routine data sets, Br. J. Cancer. 107 (2012) 1220–1226. doi:10.1038/bjc.2012.408.

[65] NHS Digital, Cervical Screening Programme England, 2017-18, (2018).

[66] NHS Digital, Breast Screening Programme England, 2016-17, 2018.

[67] Y. Hirst, S. Stoffel, G. Baio, L. McGregor, C. von Wagner, Uptake of the English Bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out, Eur. J. Cancer. 103 (2018) 267–273. doi:10.1016/j.ejca.2018.07.135.

[68] A.J. Quyn, C.G. Fraser, G. Stanners, F.A. Carey, C. Carden, A. Shaukat, R.J. Steele, Uptake trends in the Scottish Bowel Screening Programme and the influences of age, sex, and deprivation, J. Med. Screen. 25 (2018) 24–31. doi:10.1177/0969141317694065.

[69] C. von Wagner, G. Baio, R. Raine, J. Snowball, S. Morris, W. Atkin, A. Obichere, G. Handley, R.F. Logan, S. Rainbow, S. Smith, S. Halloran, J. Wardle, Inequalities in participation in an organized national colorectal cancer screening programme: results from the first 2.6 million invitations in England, Int. J. Epidemiol. 40 (2011) 712–8. doi:10.1093/ije/dyr008.

[70] C.K. Palmer, M.C. Thomas, C. von Wagner, R. Raine, Reasons for non-uptake and subsequent participation in the NHS Bowel Cancer Screening Programme: a qualitative study, Br. J. Cancer. 110 (2014) 1705–11. doi:10.1038/bjc.2014.125.

[71] L. Irwig, K. McCaffery, G. Salkeld, P. Bossuyt, Informed choice for screening: implications for evaluation, BMJ. 332 (2006) 1148–50. doi:10.1136/bmj.332.7550.1148.

[72] A. Chapple, S. Ziebland, P. Hewitson, A. McPherson, What affects the uptake of screening for bowel cancer using a faecal occult blood test (FOBt): a qualitative study, Soc. Sci. Med. 66 (2008) 2425–35. doi:10.1016/j.socscimed.2008.02.009.

[73] S.H. Lo, S. Halloran, J. Snowball, H. Seaman, J. Wardle, C. von Wagner, Colorectal cancer screening uptake over three biennial invitation rounds in the English bowel cancer screening programme, Gut. 64 (2015) 282–291. doi:10.1136/gutjnl-2013-306144.

[74] R.J.C. Steele, I. Kostourou, P. McClements, C. Watling, G. Libby, D. Weller, D.H. Brewster, R. Black, F.A. Carey, C. Fraser, Effect of repeated invitations on uptake of colorectal cancer screening using faecal occult blood testing: analysis of prevalence and incidence screening, BMJ. 341 (2010) c5531.

[75] J. Wardle, C. von Wagner, I. Kralj-Hans, S.P. Halloran, S.G. Smith, L.M. McGregor, G. Vart, R. Howe, J. Snowball, G. Handley, R.F. Logan, S. Rainbow, S. Smith, M.C. Thomas, N. Counsell, S. Morris, S.W. Duffy, A. Hackshaw, S. Moss, W. Atkin, R. Raine, Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the English NHS Bowel Cancer Screening Programme (ASCEND): four cluster-randomised controlled trials, Lancet (London, England). 387 (2016) 751–9. doi:10.1016/S0140-6736(15)01154-X.

[76] R. Raine, S.M. Moss, C. von Wagner, W. Atkin, I.K. Hans, R. Howe, F. Solmi, S. Morris, N. Counsell, A. Hackshaw, S. Halloran, G. Handley, R.F. Logan, S. Rainbow, S. Smith, J. Snowball, H. Seaman, M. Thomas, S.G. Smith, L.M. McGregor, G. Vart, J. Wardle, S.W. Duffy, A national cluster-randomised controlled trial to examine the effect of enhanced reminders on the socioeconomic gradient in uptake in bowel cancer screening, Br. J. Cancer. 115 (2016) 1479–1486. doi:10.1038/bjc.2016.365.

[77] P. Hewitson, A.M. Ward, C. Heneghan, S.P. Halloran, D. Mant, Primary care endorsement letter and a patient leaflet to improve participation in colorectal cancer screening: results of a factorial randomised trial, Br. J. Cancer. 105 (2011) 475–80. doi:10.1038/bjc.2011.255.

[78] J. Shankleman, N.J. Massat, L. Khagram, S. Ariyanayagam, A. Garner, S. Khatoon, S. Rainbow, S. Rangrez, Z. Colorado, W. Hu, D. Parmar, S.W. Duffy, Evaluation of a service intervention to improve awareness and uptake of bowel cancer screening in ethnically-diverse areas, Br. J. Cancer. 111 (2014) 1440–7. doi:10.1038/bjc.2014.363.

[79] J. Digby, P.J. McDonald, J.A. Strachan, G. Libby, R.J. Steele, C.G. Fraser, Use of a faecal immunochemical test narrows current gaps in uptake for sex, age and deprivation in a bowel cancer screening programme, J. Med. Screen. 20 (2013) 80–85. doi:10.1177/0969141313497197.

[80] S. Moss, C. Mathews, T.J. Day, S. Smith, H.E. Seaman, J. Snowball, S.P. Halloran, Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England, Gut. 66 (2017) 1631–1644. doi:10.1136/gutjnl-2015-310691.

[81] UK National Screening Committee, Note of the meeting held on the 19 November 2015 at Goodenough College, London, 2015.

[82] C.G. Fraser, C.M. Matthew, N.A.G. Mowat, J.A. Wilson, F.A. Carey, R.J.C. Steele, Immunochemical testing of individuals positive for guaiac faecal occult blood test in a screening programme for colorectal cancer: an observational study, Lancet. Oncol. 7 (2006) 127–31. doi:10.1016/S1470-2045(05)70473-3.

[83] J. Digby, C.G. Fraser, F.A. Carey, P.J. McDonald, J.A. Strachan, R.H. Diament, M. Balsitis, R.J.C. Steele, Faecal haemoglobin concentration is related to severity of colorectal neoplasia, J. Clin. Pathol. 66 (2013) 415–419. doi:10.1136/jclinpath-2013-201445.

[84] J.M. Auge, M. Pellise, J.M. Escudero, C. Hernandez, M. Andreu, J. Grau, A. Buron, M. López-Cerón, X. Bessa, A. Serradesanferm, M. Piracés, F. Macià, R. Guayta, X. Filella, R. Molina, W. Jimenez, A. Castells, Risk Stratification for Advanced Colorectal Neoplasia According to Fecal Hemoglobin Concentration in a Colorectal Cancer Screening Program, Gastroenterology. 147 (2014) 628–636.e1. doi:10.1053/j.gastro.2014.06.008.

[85] I. Stegeman, T.R. de Wijkerslooth, E.M. Stoop, M.E. van Leerdam, E. Dekker, M. van Ballegooijen, E.J. Kuipers, P. Fockens, R.A. Kraaijenhagen, P.M. Bossuyt, Combining risk factors with faecal immunochemical test outcome for selecting CRC screenees for colonoscopy, Gut. 63 (2014) 466–71. doi:10.1136/gutjnl-2013-305013.

[86] J.A. Cooper, N. Parsons, C. Stinton, C. Mathews, S. Smith, S.P. Halloran, S. Moss, S. Taylor-Phillips, Risk-adjusted colorectal cancer screening using the FIT and routine screening data: development of a risk prediction model, Br. J. Cancer. 118 (2018) 285–293. doi:10.1038/bjc.2017.375.

[87] W.S. Atkin, J. Cuzick, J.M. Northover, D.K. Whynes, Prevention of colorectal cancer by once-only sigmoidoscopy, Lancet (London, England). 341 (1993) 736–40.

[88] W.S. Atkin, B.C. Morson, J. Cuzick, Long-Term Risk of Colorectal Cancer after Excision of Rectosigmoid Adenomas, N. Engl. J. Med. 326 (1992) 658–662. doi:10.1056/NEJM199203053261002.

[89] P. Schoenfeld, S. Lipscomb, J. Crook, J. Dominguez, J. Butler, L. Holmes, D. Cruess, D. Rex, Accuracy of polyp detection by gastroenterologists and nurse endoscopists during flexible sigmoidoscopy: a randomized trial, Gastroenterology. 117 (1999) 312–8.

[90] W.S. Atkin, C.F. Cook, J. Cuzick, R. Edwards, J.M.A. Northover, J. Wardle, Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial, Lancet. 359 (2002) 1291–300. doi:10.1016/S0140-6736(02)08268-5.

[91] N. Segnan, P. Armaroli, L. Bonelli, M. Risio, S. Sciallero, M. Zappa, B. Andreoni, A. Arrigoni, L. Bisanti, C. Casella, C. Crosta, F. Falcini, F. Ferrero, A. Giacomin, O. Giuliani, A. Santarelli, C.B. Visioli, R. Zanetti, W.S. Atkin, C. Senore, SCORE Working Group, Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of

the Italian Randomized Controlled Trial—SCORE, J. Natl. Cancer Inst. 103 (2011) 1310–22. doi:10.1093/jnci/djr284.

[92] UK National Screening Committee, Note of the meeting held on 10 March 2011 at Royal Free Hampstead NHS Trust, 2011.

Chapter 3. Colorectal Cancer Ascertainment Through Cancer Registries, Hospital Episode Statistics, And Self-Reporting Compared to Confirmation by Clinician

A Cohort Study Nested Within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

The current chapter presents in manuscript-form work that has undergone peer review [1]. The study design was conceived with help from Professor Usha Menon, Dr Aleksandra Gentry-Maharaj, Dr Any Ryan, and Dr Matthew Burnell. I personally undertook all analyses, including the creation of all Figures and Tables, and wrote all drafts with advice from Dr Sophia Apostolidou, Dr Wendy Alderton, Dr Julie Barnes, and Dr John Timms. Dr Evangelia-Ourania Fourkala, Dr Sophia Apostolidou, Dr Andy Ryan, and Dr Aleksandra Gentry-Maharaj was involved in obtaining clinician's confirmation.

3.1 Background

Electronic health records (EHRs) are datasets created for routine administrative purposes. Increasingly, however, they are used to assess health outcomes in large observational studies and randomised controlled trials [2–6]. Evaluating their quality is therefore crucial.

National registries are responsible for cancer registration (CR) in Northern Ireland, Scotland & Wales, while one of eight regional hubs register cancers diagnosed in England. English registries have been shown by one-directional comparisons to capture 98% of CRCs recorded in routine healthcare databases [7,8]. However, such estimates are likely overestimated since it is unlikely for an individual database to have complete coverage of all events, while iterative refinements to the registration process require ongoing evaluations.

Hospital Episode Statistics (HES) is an administrative dataset that documents all admissions and attendances to NHS Trusts in England. Its secondary uses for research has been reviewed [9]. HES is appealing since the compulsory recording of hospital events amasses ~125-million diagnostic and procedural records amenable to digital analyses per annum [10], while their coding from case notes by dedicated professionals is generally

accurate [11,12]. Despite the comprehensive coverage of CRC events HES records [13], its reliability has not been determined.

Self-reporting (SR), meanwhile, remains an option for identifying events where EHRs are unavailable or inaccessible. However, the reliability of self-reporting is currently unclear, particularly despite an increased detection of adenomatous polyps [14] and potential confusion resulting from repeat testing due to technical issues and false positive screening tests [15].

We herein determined the feasibility of using electronic health records and self-reporting for colorectal cancer (CRC) ascertainment in UK women. We report on the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of cancer & death registrations, HES, and self-reporting relative to clinician's confirmation obtained from the patients' clinician. We also explore the reliability of self-reporting of CRC and the factors that determine the accurate self-reporting of CRC.

3.2 Methods

Study design. The present retrospective cohort study was nested within the UK Collaborative Trial of Ovarian Screening. During 2001–05, 202,638 postmenopausal women aged 50–74 were randomised to the UKCTOCS trial centres across England, Northern Ireland & Wales following an invitation from Health Authorities [16]. All were asked to provide their ethnicity, postcode, height, and weight by postal questionnaire at recruitment. Postcodes were updated throughout the trial. The study cohort were women who were identified as having been diagnosed with CRC since randomisation (see *Data sources*) and gave consent for access to their medical records (see *Clinician's confirmation*).

Data sources

Electronic health records. Women were linked to their EHR by NHS number. CRs were up-to-date until 19 May 2011 for England & Wales and 23 February 2010 for Northern Ireland. DCs were up-to-date until 19 May 2011 for England & Wales and 9 March 2011 for Northern Ireland. HES in and outpatient episodes were received for those in England from 2001 and 2003, respectively, until 22 July 2010. CR, DC, and HES records were reviewed for malignant neoplasms of the colon or rectum (ICD10 C18–C20) diagnosed following randomisation to the trial. Duration of follow-up was calculated from the date of randomisation to the date of the latest update (or date of loss to follow-up if before).

Postal questionnaire. All participants were able to self-report CRC using postal questionnaires first sent at 3–5-years post-randomisation (FUQ I (Supplementary Figure S2)) and thereafter in April 2014 (FUQ II (Supplementary Figure S3)) [16]. Each questionnaire included an item on whether the woman had been diagnosed with 'bowel/colorectal cancer'. FUQ I also requested data on their education (college/university degree or nursing/teaching qualification or A-/O-level or vocational qualification or other or none of the above), and alcohol (how many units consumed on average per week) & tobacco (ever/never) use.

Trial database. Incidental notifications were received via the UKCTOCS staff due to 1) investigation of a possible ovarian cancer diagnosis after a positive screening result or 2) a participant citing colorectal cancer diagnosis as a reason for withdrawal from the trial. Half of all UKCTOCS women were screened via annual serum CA125 (25%) or transvaginal ultrasound (25%), while the remaining half were controls who received no screening.

Clinician's confirmation. All women for whom a CRC notification was received from any of the above sources by 24 May 2011 were identified. In May 2012, a postal questionnaire (CRCQ (Supplementary Figure S4)) was sent to the treating clinician (GP by default or treating hospital consultant if self-reported in the FUQ I), requesting the diagnosis date, primary site, stage, grade, morphology, and treatment details (Supplementary Figure S4). The CRCQ requested details of a specialist if the initial contact was unable to provide complete data, who was contacted if necessary. Reminder CRCQs were sent to nonresponding clinicians. Questionnaire outcomes (confirmed CRC, benign polyp, no CRC or benign polyp) were captured in the UKCTOCS Trial Management System. Where multiple CRCQs per women were obtained, a confirmed CRC superseded one which reported a benign diagnosis. The earliest of two cancer diagnosis dates was used where necessary. Cancers without a diagnosis date and those diagnosed after 24 May 2011 were excluded.

Data analyses

Non-response bias. Likelihood of a non-response bias was accessed by multivariable analysis (Kruskal-Wallis/Fisher's Exact & post hoc pairwise Wilcox rank sum/$\chi 2$ with

Bonferroni adjustment) of the characteristics and composition of notifications for women with verified TP & FP and those whose clinician did not respond.

Sensitivity & PPV. The number of true positive (TP), true negative (TN), false positive (FP), and false negatives (FN) notifications for each dataset were calculated by pairwise comparison to clinician's confirmation (gold-standard). The sensitivity (TP/(TP+FN)), PPV (TP/(TP+FP)), and 95% confidence intervals for each notification source were then computed. The CR analysis included events diagnosed ≥ one year before the latest registry update (19 May 2010 for England & Wales and 23 February 2009 for Northern Ireland). The DC analysis included only cancers with death dates before the latest update (19 May 2011 for England & Wales or 9 March 2011 for Northern Ireland). For HES, only women from English centres with a diagnosis date before 22 July 2010 (HES update 22 July 2010) were included. Our SR analysis included women who had a diagnosis/notification date before the date their FUQ I or II was returned. The use of EHR/SR to complement delays in CRs was also assessed via comparison of the adjunct sensitivities and PPVs for CRCs diagnosed ≥ one year in advance of latest registry update. Adjunct analyses of CR & HES and CR & DC & HES were restricted to England.

Timeliness and completeness. We assessed the completeness of CRs by determining the proportion (%) of all confirmed cancers that were registered in updates received by 4 September 2016 for England & Wales and 15 April 2015 for Northern Ireland. Curation times for CRs were defined as the years between diagnosis and latest update.

Specificity & NPV. Specificities and NPVs were estimated relative to the expected number of cases derived from an age-standardised rate of 57.2 cases per 100,000 person-years [17] applied to the total years of follow-up for subsets of all 202,365 consenting women who 1) lived in the UK (CR analysis), 2) had died before the latest DC update (DC analysis), 3) lived in England (HES analysis), and 4) who returned their FUQ before 24 May 2011 (SR analysis). Total years of follow-up was the sum of all years followed-up from randomisation to latest update (see *Electronic health records*), or date of death if before. Column totals of the confusion matrix were inferred from the expected number of cases (TP+FN) or sample size minus expected cases (FP+TN). Row totals of the confusion matrix were inferred from the number of notifications (TP+FP) or sample size minus number of notifications (FN+TN). Using the sensitivity estimates (Table 3–3), the

number of TPs is as 1) sensitivity * (TP+FN). Individual cells (i.e. FN, TP & TN) were then deduced from subtraction of TP from column/row totals. The specificity (TN/(FP+TN)), NPV (TN/(FN+TN)), and 95% confidence intervals were then computed. Proofs are presented in Supplementary Method M1.

Determinants of accurate self-reporting. A binomial logistic regression model was fitted to identify the variables predictive for self-reporting concordant with the clinician's confirmation. The analysis was restricted to those who returned their FUQ. Outcomes were either concordant (TP & TN) or discordant (FP & FN) self-reporting. Predictive determinants were age (at self-report), BMI (at recruitment to the UKCTOCS), education (low: A-/O-level or vocational qualification; high: college/university degree or nursing/teaching qualification; other: none of the above), alcohol (non-drinker/< 1 unit per day/≥ 1 unit per day), tobacco (never/ever), and socioeconomic status (Index of Multiple Deprivation 2015 (IMD) derived from postcode (England only)). The IMD 2015 scoring is a composite measure of seven socioeconomic indicators that stratifies all English postcodes on a gradient from most to least deprived [18]. Outliers (> 3 SD of the mean/± 1.5 IQR) were entered as missing. Missing data were imputed five times using predictive mean matching. Logit coefficients for each imputed dataset were pooled according to Rubin's rules [19]. Odds ratios were adjusted for all variables listed.

Statistical analyses. Five-way Venn diagrams were produced using the online InteractiVenn tool [20]. All other analyses were made with R version 3.3.2 [21] running the epiR [22], ggplot2 [23], and mice [24] packages. A p value < 0.05 was considered significant**.**

3.3 Results

We received 2,217 notifications of a post-randomisation CRC diagnosis for 1,085 women (Figure 3–1). These included 814 CRs, 233 DCs, 625 HES, 400 SRs (FUQ I only), and 145 notifications via the trial database. Clinicians of all 1,085 women were sent a CRCQ. Responses were received from 660 (61%). Nineteen women were excluded as 18 of the confirmed cancers had no diagnosis date on the returned CRCQ and one cancer was diagnosed after May 2011 (self-reported in 2006). Overall, 641 women were eligible for analysis. Of these, 514 had a verified CRC, 24 had a benign polyp, while 103 had never

been diagnosed with malignant or benign colorectal disease. Histology reports were disclosed for 15% (75/514) of verified CRCs.

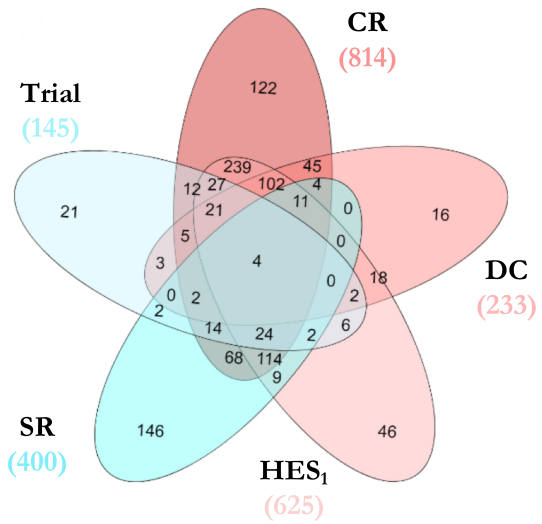


Figure 3–1: CRC notifications by source. CRC notifications by source. 1 England only. Abbreviations: CR, cancer registration; DC, death certification; HES, Hospital Episode Statistics; SR, self-reporting.

There were 1,173 TP notifications and 152 FP notifications (38 for benign disease & 114 for no malignant/benign diagnosis). The contribution of each notification to the pools of TPs and FPs is presented in Figure 3–2. Cancer registrations contributed to 38.6% (453) of 1,173 TP notifications, followed by HES (27.9% (327)), self-reporting (18.0% (211)), death certificates (8.4% (98)), and trial resources (7.2% (84)). Cancer registrations provided the greatest contribution of uniquely recorded TPs (4.6% (54)), while self-reporting (0.7% (8)) and death certificates (0.3% (4)) provided few. Overlap in TP events between two datasets ranked in order of CR & HES (38.2% (298/780)), CR & SR (29.1% (193/664)), HES & SR (23.8% (128/538)), DC & HES (16.2% (69/425)), CR & DC (15.8% (87/551)), and DC & SR (2.6% (8/309)) (Figure 3–2a). Contributions to the pool of 152 FP notifications ranked in order of self-reporting (62.5% (95)), cancer registrations (16.4% (25)), HES (14.5% (22)), trial resources (5.3% (8)), and death certificates (1.3% (2)) (Figure 3–2b). Notably, 58.6% (89) of FPs were uniquely notified by self-reporting—8 where a benign adenoma was present and 84 where no benign/malignant diagnoses were verified.

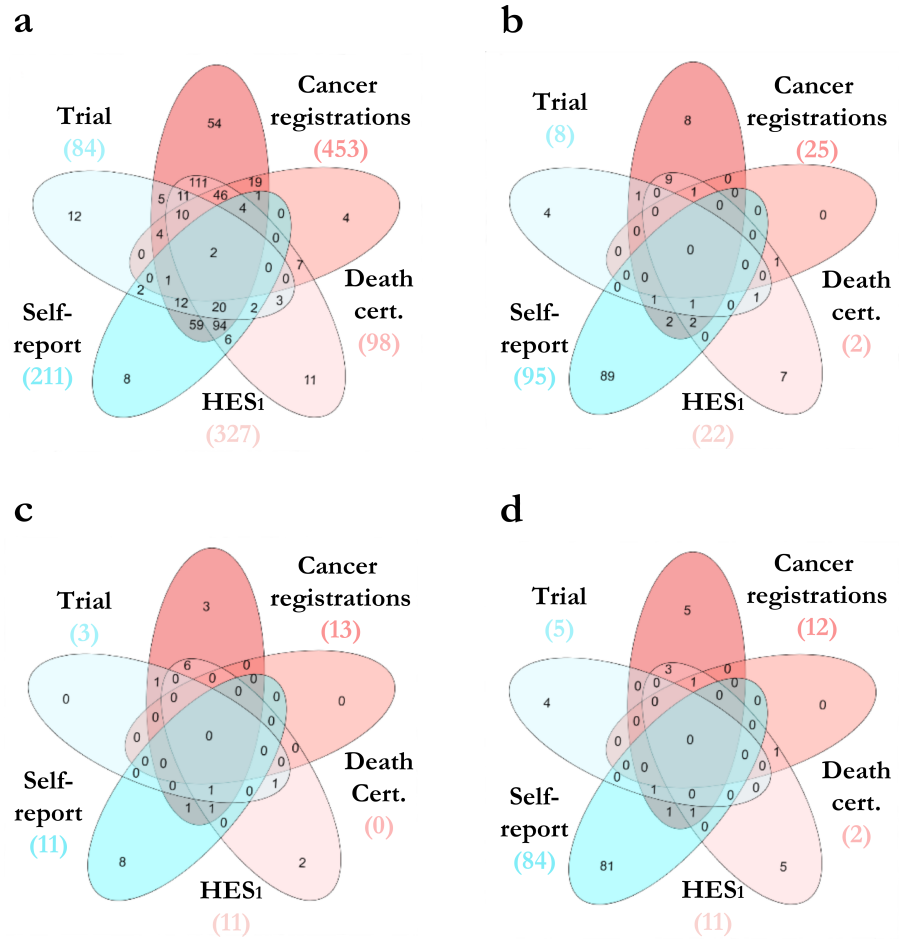


Figure 3–2: Contribution of follow-up routes to the true (a) and false positive notifications (b), stratified to those with benign polyp (c) and no diagnosis of benign or CRC (d). 1 England only. Abbreviations: HES, Hospital Episode Statistics.

Baseline characteristics are presented in Table 3–1. Women whose clinicians did not respond (NR) had a greater proportion of notifications informed by DC (15%) to TPs (8%) and FPs (1%). NRs and TPs were similar but differed to FPs in their proportion of notifications informed by CR (37, 39%, & 16%, respectively), HES (26, 28, and 14%), and SR (16, 18, & 62%). NRs were comparable to TPs in their age, BMI, ethnicity, education, and alcohol use and smoking status and differed to FPs in age, IMD score, education, and alcohol use and smoking status (Table 3–1). NRs had a markedly greater proportion of deaths (38%) than FPs (5%) and TPs (24%).

Table 3–1: Non-responders show notable similarities to True Positives and differences to False Positives.

	Responders (TP)	Responders (FP)	Non-responders	*p* value[1]
		Median (Range)		
Age	72 (57–83)	68 (57–82)	72 (57–83)	NR vs. FP (0.018)
BMI (Kg m-2)	25.8 (10.3–110.8)	26.4 (18.7–150.0)	25.6 (0.1–47.2)	0.753
IMD score	13.6 (1.6–73.9)	11.4 (2.5–74.4)	16.4 (1.6–70.8)	NR vs. TP (0.011) & FP (0.003)
		Count (%)		
Cohort size	**514 (100)**	**127 (100)**	**425 (100)**	
Alcohol				NR vs. FP (< 0.001)
Non-drinker	83 (16)	16 (13)	71 (17)	
< 1 unit a day	178 (35)	67 (53)	124 (29)	
≥ 1 unit a day	71 (14)	37 (29)	55 (13)	
Missing	182 (35)	7 (6)	175 (41)	
Deaths[2]				NR Vs. TP (< 0.001) & FP (< 0.001)
Alive	390 (76)	121 (95)	263 (62)	
Deceased	124 (24)	6 (5)	162 (38)	
Education				NR vs. FP (< 0.001)
Low	131 (25)	50 (39)	104 (24)	
High	89 (17)	43 (34)	54 (13)	
Other	106 (21)	25 (20)	95 (22)	
Missing	188 (37)	9 (7)	172 (40)	
Ethnicity				0.821
White	493 (96)	124 (98)	411 (97)	
Black	9 (2)	2 (2)	7 (2)	
Other	9 (2)	1 (1)	3 (1)	
Missing	3 (1)	0 (0)	4 (1)	
Tobacco				NR vs. FP (< 0.001)
Ever	109 (21)	43 (34)	120 (28)	
Never	187 (36)	64 (50)	135 (32)	
Missing	218 (42)	20 (16)	170 (40)	
Notifications	**1,173 (100)**	**152 (100)**	**841 (100)**	NR vs. TP (0.001) & FP (< 0.001)
CR	453 (39)	25 (16)	312 (37)	
DC	98 (8)	2 (1)	125 (15)	
HES	327 (28)	22 (14)	220 (26)	
SR	211 (18)	95 (62)	137 (16)	
Trial	84 (7)	8 (5)	47 (6)	

[1] *p* values refer to the multivariable test (Kruskal-Wallis/Fisher's Exact) if no significance detected or the Bonferroni-adjusted, pairwise post hoc test(s) (pairwise Wilcox rank sum/χ^2) if significance detected. [2] at clinician's confirmation. Abbreviations: TP; True Positives (Responders); FP, False Positives (Responders); NR, Non-responders; CRCQ, Colorectal Cancer Questionnaire; CR, Cancer Registry; DC, Death Certificate; HES, Hospital Episode Statistics; SR, Self-report.

Cancer registrations. Sensitivity and PPV were estimated from 618 verified women (491 with CRC diagnosed one year before the latest update, 24 with a benign polyp, and 103 with no colorectal disease). Median follow-up from randomisation to registry update or loss to follow-up was 6.5 years (IQR 2.1; n 618). Curation times ranged from 1.0–9.1 years (median 4.1; IQR 3.2; n 491).

Table 3–2: True and false notifications by data source.

a

CR[1]	CC +	CC –
+	453 TP	25 FP
–	38 FN	102 TN

b

CR[2]	CC +	CC –
+	485 TP	23 FP
-	6 FN	104 TN

c

DC	CC +	CC –
+	98 TP	2 FP
–	3 FN	0 TN

d

HES	CC +	CC –
+	327 TP	22 FP
-	70 FN	83 TN

e

SR[3]	CC +	CC –
+	211 TP	95 FP
–	303 FN	32 TN

f

SR[4]	CC +	CC –
+	211 TP	95 FP
-	22 FN	25 TN

[1] 1–9 years curation. [2] 6–14 years curation. [3] non-responses negative by default. 4 responders only. Abbreviations: CC, clinician's confirmation; CR, cancer registration; DC, death certification; HES, Hospital Episode Statics; SR, self-reporting; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

CR notified of 54 unique TP events not captured by other sources. There were 38 CRCs without a CR after a minimum of one-year curation. Of these, 32 were registered when curation was extended to 6.3–14.4 years (median 9.4; IQR 3.2), while six residents in England remained unregistered after 6.8, 7.5, 10.5, 11.5, 13.1, and 13.6 years of curation (Table 3–2b & Table 3–4). Four of the FP registrations (2 benign and 2 no CRC) were rescinded by 2015/16, while 2 TN (1 benign and 1 no CRC) became FP. Overall, a further

36 women were correctly classified (32 CRC TPs and 2 benign & 2 no CRC TNs), and two were incorrectly registered (1 benign & 1 no CRC FPs) when curation was extended.

The sensitivity and PPV of CR after allowing a minimum of 1 year for their curation were 92% (453/491; 95% CI 90–94) and 95% (453/478; 95% CI 92–97), respectively (Table 3–3). When supplemented with registrations received until 2015/16, the sensitivity of CR increased to 99% (485/491; 95% CI 97–100), while the PPV remained at 95% (485/508; 95% CI 93–97). Sensitivities by year of diagnosis (2002–2010 inclusively) after a minimum of one or six years of curation are presented in Figure 3–3 . Specificity and NPV were estimated from 202,230 women. Relative to the 850 cases expected after 1,486,350 person-years follow-up, the specificity and NPV were 100% (201,485/201,515; 95% CI 100–100) and 100% (201,485/201,551; 95% CI 100–100), respectively.

Table 3–3: Performance estimates for Electronic Health Records and self-reporting.

Dataset	Sensitivity[1] (95% CIs)	PPV[1] (95% CIs)	Specificity[2] (95% CIs)	NPV[2] (95% CIs)
CR[3]	0.92 (0.90–0.94)	0.95 (0.92–0.97)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
CR[4]	0.99 (0.97–1.00)	0.95 (0.93–0.97)	—	—
DC	0.97 (0.92–0.99)	0.98 (0.93–1.00)	0.97 (0.97–0.97)	1.00 (1.00–1.00)
HES	0.82 (0.78–0.86)	0.94 (0.91–0.96)	1.00 (1.00–1.00)	1.00 (1.00–1.00)
SR[4]	0.91 (0.86–0.94)	0.69 (0.63–0.74)	1.00 (1.00–1.00)	1.00 (1.00–1.00)

[1] relative to clinician's confirmation. [2] relative to expected cases. [3] 1–9 years curation (median 4.1, IQR 3.2). [4] 6–14 years curation (median 9.4, IQR 3.2). Abbreviations: CIs, confidence intervals; PPV, positive predictive value; CR, cancer registration; DC, death certificate; HES, Hospital Episode Statistics; SR, self-reporting.

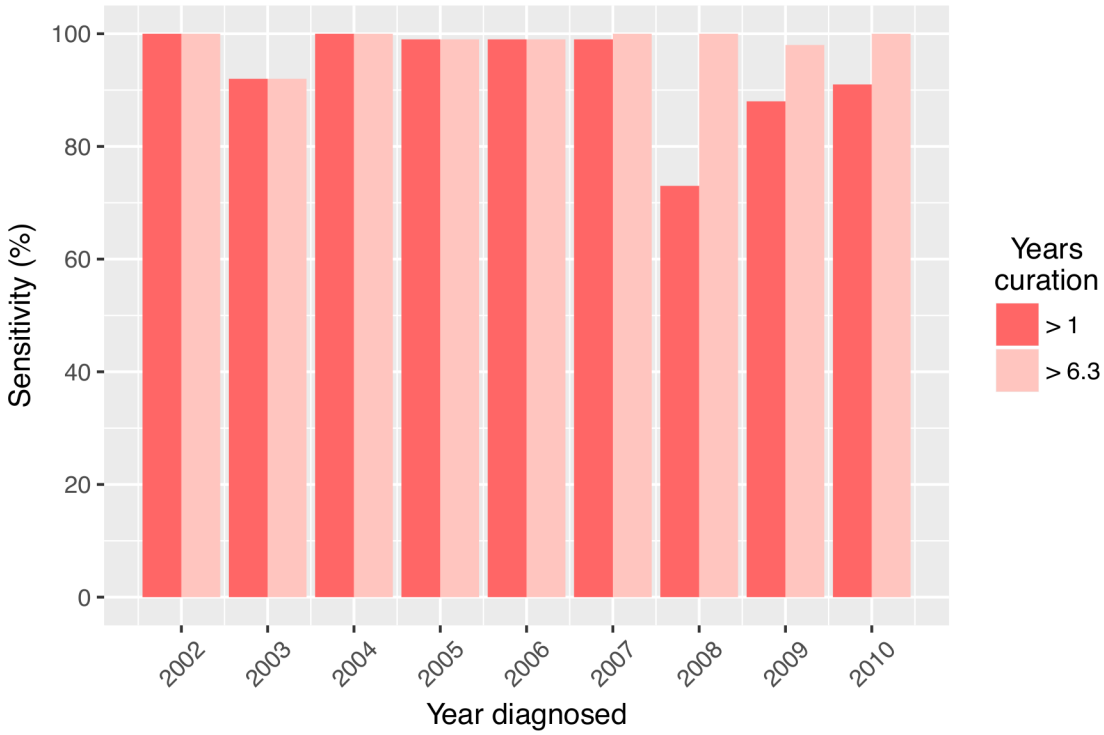


Figure 3–3: Sensitivity of cancer registrations curated after a minimum of 1 & 6.3 years. Sensitivities by year of diagnosis (2002–10 inclusively) were 100, 92, 100, 99, 99, 99, 73, 88, and 91% with allowance of 1.0–9.1 years curation (CR update in 2010/11) and 100, 92, 100, 99, 99, 100, 100, 98, and 100% when supplemented with registrations received until 2015/16 (6.3–14.4 years curation).

Table 3–4: Longitudinal evolution of cancer registrations held for the same women.

Verified CRC			Benign polyp			No CRC		
2010–11	**2015–16**	**Count**	**2010–11**	**2015–16**	**Count**	**2010–11**	**2015–16**	**Count**
+	+	453	+	+	11	+	+	10
–	+	32	–	+	1	–	+	1
+	–	0	+	–	2	+	–	2
–	–	6	–	–	10	–	–	90

+ Cancer registration. – No cancer registration. Thirty-two of the women who had a confirmed diagnosis of CRC and who had no cancer registration in 2010–11 were registered by 2015–16. Only six of the 491 confirmed diagnoses had no cancer registration by 2015–16. Of the 25 false-positive registrations held for women who had either a diagnosis of a benign polyp or had no diagnoses at all, four were rescinded, and 21 remained in 2016. Two women who were true negatives were subsequently registered by 2016.

Death Registrations. Sensitivity and PPV were estimated from the 104/641 (16.2%) verified women who died before the latest DC update (102 cancers diagnosed before the latest update, 0 benign, and 2 with no cancer/benign (Table 3–2). Median follow-up from randomisation to registry update or loss to follow-up was 4.8 years (IQR 3.2; n 104). The sensitivity and PPV of DCs were 97% (98/102; 95% CI 92–99) and 98% (98/101; 95% CI 93–100), respectively (Table 3–3). Specificity and NPV were estimated from 7,202 registered deaths. Relative to the 19 cases expected after 33,968 person-years follow-up, the specificity and NPV were 97% (6,968/7,183; 95% CI 97–97) and 100% (6,968/6,969; 95% CI 100–100), respectively.

Hospital Episode Statistics. Sensitivity and PPV were estimated from 502 verified women who were randomised from centres in England (397 with confirmed CRC diagnosed before 22 July 2010, 21 with a benign polyp, and 84 with no colorectal disease (Table 3–2d)). Median follow-up from randomisation to HES update or loss to follow-up was 6.7 years (IQR 2.2; n 502). Eleven TP HES notifications were unique and were all diagnosed in 2008–09. The sensitivity and PPV for HES notifications were 82% (327/397; 95% CI 78–86) and 94% (327/349; 95% CI 91–96), respectively (Table 3–3). Specificity and NPV were estimated from 157,839 women in England. Relative to the 616 cases expected after 1,076,512 person-years follow-up, the specificity and NPV were 100% (157,105/157,223; 95% CI 100–100) and 100% (157,105/157,214; 95% CI 100–100), respectively.

Self-reporting. Of 641 eligible women, 353 (55.1%) completed a FUQ (291 FUQ1, 62 FUQ II). Of these, 233 had a confirmed CRC, 22 had a benign polyp, and 98 had no malignant/benign colorectal disease (Table 3–2e–f). The sensitivity and PPV of SR in the cohort of responders were 91% (211/233; 95% CI 86–94) and 69% (211/306; 95% CI 63–74), respectively (Table 3–3). The sensitivity where nonresponses were negative by default was 41% (211/514; 95% CI 37–45) at a 69% PPV. Specificity and NPV were estimated from 144,313 women who returned their FUQ. Relative to the 321 cases expected after 561,274 person-years follow-up, the specificity and NPV were 100% (143,883/143,992; 95% CI 100–100) and 100% (143,883/143,913; 95% CI 100–100), respectively.

We fitted a logistic regression to ascertain the variables associated with concordant self-reporting. The adjusted ORs are presented in Table 3–5. Missing data

were imputed for IMD (83 (9 England, 22 Northern Ireland, 52 Wales)), smoking (51), education (13), alcohol (8), age at self-report (4), and BMI (3 including 1 outlier). Age at self-report markedly increased the odds of being concordant with their clinician's confirmation (adjusted OR 1.05; 95% CI 1.01–1.10; *p* 0.026). No other associations were statistically significant.

Table 3–5: Variables associated with self-reporting concordant with clinician's confirmation.

Variable	OR [1] (95% CIs)	*p* value
(Intercept)	0.22 (0.01–7.53)	0.395
Age (year)	1.05 (1.01–1.10)	0.026 *
BMI (Kg/m²)	0.98 (0.91–1.05)	0.504
IMD score	1.01 (0.98–1.04)	0.441
Alcohol		
0	1.00	
< 1 U/day	0.48 (0.20–1.13)	0.091
> 1 U/day	0.38 (0.15–1.00)	0.051
Tobacco		
Never	1.00	
Ever	1.18 (0.63–2.19)	0.604
Education		
Low	1.00	
High	1.19 (0.60–2.36)	0.613
Other	0.86 (0.41–1.83)	0.698

[1] Adjusted for all variables listed. n 353. Age is at self-report. Abbreviations: OR, odds ratio; BMI, Body Mass Index; IMD, Index of Multiple Deprivation.

Trial database. Notification via the UKCTOCS trial centre accounted for few of the overall TP (7.2%; 84/1,173), FP benign (7.9%; 3/38), and FP no CRC/benign notifications (4.4%; 5/114) but captured 12 events that would have otherwise been missed (Figure 3–2a). Reasons for withdrawal from the UKCTOCS informed the majority of TPs (84.5% (71/84)) and FPs (87.5% (7/8)).

Adjunct datasets. Given the demonstrated delay in cancer registrations, researchers may be interested in how best to supplement their analyses. There were 38 cancers not registered 1 year after diagnosis (36 in England). HES, SR, and DC captured 77.8% (28/36), 36.8% (14/38), and 21.1% (8/38) of these events, respectively. The sensitivity and PPV of CR & HES (n 501) were 98% (388/396; 95% CI 96–99) and 92% (388/422; 95% CI 89–94); CR & SR (n 618): 95% (467/491; 95% CI 93–97) and 80% (467/581;

95% CI 77–84); and CR & DC (n 618): 94% (461/491; 95% CI 91–96) and 95% (461/487; 95% CI 92–96). The sensitivity and PPV of CR, DC & HES combined (n 501) were 98% (388/396; 95% CI 96–99) and 92% (388/422; 95% CI 89–94).

3.4 Discussion

Principle findings. Advances in healthcare are achieved through high-quality epidemiological studies informed by a comprehensive and reliable ascertainment of outcomes. We have herein evaluated the performance of EHRs and self-reporting for ascertaining diagnoses of CRC in UK women in comparison to a patient's clinical records. We found that 92% of CRCs diagnosed in the UK during 2001–10 were registered within one year, and 99% within six years. Researchers wishing to overcome delays in the curation of cancer registrations are advised to use HES in adjunct, which combined with CR at one-year curation had a sensitivity of 98% and a PPV of 92%. Finally, self-reporting of CRC for standalone or adjunct ascertainment was not useful owing to high false positivity and low response rates.

Strengths and limitations. We acknowledge several limitations. Firstly, we did not verify the absence of cancer in those without notification, and thus the reliability of sensitivity estimates is dependent on the number of false negatives missed. Secondly, the potential for bias in the 39% of non-responding clinicians contacted should not be overlooked for an underascertainment of FPs through non-response would overstate the PPV, while an underascertainment of TPs would underestimate the sensitivity. Nonetheless, barring a higher proportion of mortality at time of clinician contact and notifications via death certificate, non-responders were in closer alignment to TPs than FPs in the notifications received, and we suspect the risk of bias to be minimal. Next are issues of generalisability. Men are disproportionally affected by CRC [17], while also being less likely to undertake FOBT screening than women [25], but were not studied here. Ethnic minorities, too, are less likely to undertake screening [25] but were underrepresented in our cohort. Finally, our findings may be affected by a 'healthy-volunteer' bias that typically affects those willing to enrol in trials [26], while it would also be reasonable to assume that recall biases arising from past diagnoses of CRC in the family or prior colonoscopy for suspected cancer could aid concordant self-reporting and would, ideally, be accounted for in analyses if obtained. Weaknesses notwithstanding, our study updates the current performance estimates of colorectal cancer registrations [6,7]. Our estimates are reliable

though us studying verified events identified via multiple routes. It is strengthened further from the high rate at which patients' electronic health records were linked.

Findings in relation to existing literature. Registration of cancers by regional registries are the cornerstone of national and international cancer surveillance [27] and have hitherto informed the implementation of colorectal cancer screening programmes [2,3]. Incomplete ascertainment of events, however, can skew analyses. Nonetheless, our group previously concluded that incomplete registration of CRC is unlikely as 85% of self-reported (but not verified) CRCs were registered within five years of diagnosis [28]. The higher sensitivity of CRC cancer registration reported here (92% vs 85%) is likely due to our exclusion of previously unknown false positive SR CRCs. Elsewhere, it is reported that 98% of surgically treated CRCs recorded in HES during 2001–07 were registered [7], while 98% of primary care records were captured within four years [8].

We found no issues with the reliability of HES (PPV 94%) in alignment with a recent meta-analysis [11]. HES did, however, have limited sensitivity that would likely preclude its use for standalone ascertainment. These missed cases are likely due to coding errors, emergency admissions, death certificate-only registrations, or privately-treated cancers. Nonetheless, a similar sensitivity (83%) was reported for prostate cancers recorded by HES relative to medical notes [29].

Self-reporting of CRC by post was unreliable. While the sensitivity (91%) of responders was similar to CR at one year (92%), a low response rate (55%) limited the actual sensitivity to 41%. Furthermore, the PPV was 69%, and some women misreported their benign polyp as cancer. Studies in the USA have reported similar sensitivities (83–85%) for postal SR, albeit at a lower PPV (54%) [30,31]. Our findings align with reports that British interviewees markedly under-report CRC diagnoses in close relatives [32], but contrasts in that they reliably self-report their participation in CRC screening [33]. As screening becomes widely accepted, an ongoing evaluation of how it might influence the layperson's ability to reliably self-report would be insightful.

Implications for clinicians and policymakers. Electronic health records generated by primary, secondary, and social care are frequently used to expedite healthcare research, and, by 2020, UK patients' records spanning the spectrum of healthcare will be linked in a move recognised by parliament as being beneficial to research [34]. Our study provides evidence to support the use of EHRs in future epidemiological research.

3.5 Conclusions

EHRs in England, Northern Ireland, and Wales are a reliable resource for ascertaining CRC events in women. Researchers looking to supplement delays in the registration of CRCs should use Hospital Episode Statistics in adjunct. Self-reporting of CRC by women is neither reliable nor is it useful in adjunct with electronic health records.

3.6 References

[1] D.S. Thomas, A. Gentry-Maharaj, A. Ryan, E.-O. Fourkala, S. Apostolidou, M. Burnell, W. Alderton, J. Barnes, J.F. Timms, U. Menon, Colorectal cancer ascertainment through cancer registries, hospital episode statistics, and self-reporting compared to confirmation by clinician: A cohort study nested within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), Cancer Epidemiol. 58 (2019) 167–174. doi:10.1016/J.CANEP.2018.11.011.

[2] J.H. Scholefield, S. Moss, F. Sufi, C.M. Mangham, J.D. Hardcastle, Effect of faecal occult blood screening on mortality from colorectal cancer: results from a randomised controlled trial, Gut. 50 (2002) 840–4.

[3] W.S. Atkin, R. Edwards, I. Kralj-Hans, K. Wooldrage, A.R. Hart, J.M. Northover, D.M. Parkin, J. Wardle, S.W. Duffy, J. Cuzick, Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial, Lancet. 375 (2010) 1624–1633. doi:10.1016/S0140-6736(10)60551-X.

[4] E.M. Burns, A. Bottle, P. Aylin, A. Darzi, R.J. Nicholls, O. Faiz, Variation in reoperation after colorectal surgery in England as an indicator of surgical performance: retrospective analysis of Hospital Episode Statistics, BMJ. 343 (2011) d4836–d4836. doi:10.1136/bmj.d4836.

[5] D. Wallace, K. Walker, A. Kuryba, P. Finan, N. Scott, J. van der Meulen, Identifying patients at risk of emergency admission for colorectal cancer, Br. J. Cancer. 111 (2014) 577–580. doi:10.1038/bjc.2014.300.

[6] C. Renzi, G. Lyratzopoulos, T. Card, T.P.C. Chu, U. Macleod, B. Rachet, Do colorectal cancer patients diagnosed as an emergency differ from non-emergency patients in their consultation patterns and symptoms? A longitudinal data-linkage study in England, Br. J. Cancer. 115 (2016) 866–875. doi:10.1038/bjc.2016.250.

[7] H. Møller, S. Richards, N. Hanchett, S.P. Riaz, M. Lüchtenborg, L. Holmberg, D. Robinson, Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates, Br. J. Cancer. 105 (2011) 170–176. doi:10.1038/bjc.2011.168.

[8] A. Dregan, H. Møller, T. Murray-Thomas, M.C. Gulliford, Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study, Cancer Epidemiol. 36 (2012) 425–429. doi:10.1016/j.canep.2012.05.013.

[9] S. Sinha, G. Peach, J.D. Poloniecki, M.M. Thompson, P.J. Holt, Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomes—systematic review and recommendations for reporting, Eur. J. Public Health. 23 (2013) 86–92. doi:10.1093/eurpub/cks046.

[10] NHS Digital, Hospital Episode Statistics, (n.d.). http://content.digital.nhs.uk/hes (accessed September 11, 2017).

[11] E.M. Burns, E. Rigby, R. Mamidanna, A. Bottle, P. Aylin, P. Ziprin, O.D. Faiz, Systematic review of discharge coding accuracy, J. Public Health (Bangkok). 34 (2012) 138–148. doi:10.1093/pubmed/fdr054.

[12] Department of Health, Payment by Results Data Assurance Framework: key findings from the 2012/13 programme, 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/262027/pbr_data_assurance_framewrk_key_find_2012-13.pdf.

[13] M. Garout, H.S. Tilney, P.P. Tekkis, P. Aylin, Comparison of administrative data with the Association of Coloproctology of Great Britain and Ireland (ACPGBI) colorectal cancer database, Int. J. Colorectal Dis. 23 (2008) 155–163. doi:10.1007/s00384-007-0390-z.

[14] J.H. Scholefield, S.M. Moss, C.M. Mangham, D.K. Whynes, J.D. Hardcastle, Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up, Gut. 61 (2012) 1036–40. doi:10.1136/gutjnl-2011-300774.

[15] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut. 61 (2012) 1439–46. doi:10.1136/gutjnl-2011-300843.

[16] I.J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J.K. Kalsi, N.N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, D.N. Crump, S.K. Davies, A. Dawnay, S. Dobbs, G. Fletcher, J. Ford, K. Godfrey, R. Gunu, M. Habib, R. Hallett, J. Herod, H. Jenkins, C. Karpinskyj, S. Leeson, S.J. Lewis, W.R. Liston, A. Lopes, T. Mould, J. Murdoch, D. Oram, D.J. Rabideau, K. Reynolds, I. Scott, M.W. Seif, A. Sharma, N. Singh, J. Taylor, F. Warburton, M. Widschwendter, K. Williamson, R. Woolas, L. Fallowfield, A.J. McGuire, S. Campbell, M. Parmar, S.J. Skates, Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial, Lancet. 387 (2016) 945–956. doi:10.1016/S0140-6736(15)01224-6.

[17] Cancer Research UK, Bowel cancer incidence statistics [UK, 2015, ICD-10 C18–C20], (2018). http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence (accessed February 28, 2018).

[18] Department for Communities and Local Government, The English Indices of Deprivation 2015, 2015.

[19] J. Barnard, D. Rubin, Small-sample degrees of freedom with multiple imputation, Biometrika. 86 (1999) 948–955. doi:10.1093/biomet/86.4.948.

[20] H. Heberle, G.V. Meirelles, F.R. da Silva, G.P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams, BMC Bioinformatics. 16 (2015) 169. doi:10.1186/s12859-015-0611-3.

[21] R Core Team, R: a language and environment for statistical computing, (2016). https://www.r-project.org/.

[22] M. Stevenson, T. Nunes, C. Heuer, J. Marshall, J. Sanchez, R. Thornton, J. Reiczigel, J. Robison-Cox, P. Sebastiani, P. Solymos, K. Yoshida, G. Jones, S. Pirikahu, S. Firestone, R. Kyle, J. Popp, M. Jay, epiR: tools for the Analysis of Epidemiological Data, (2017). https://cran.r-project.org/package=epiR.

[23] H. Wickham, ggplot2: elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.

[24] S. van Buuren, K. Groothuis-Oudshoorn, mice: multivariate Imputation by Chained Equations in R, J. Stat. Softw. 45 (2011) 1–67.

[25] C. von Wagner, G. Baio, R. Raine, J. Snowball, S. Morris, W. Atkin, A. Obichere, G. Handley, R.F. Logan, S. Rainbow, S. Smith, S. Halloran, J. Wardle, Inequalities in participation in an organized national colorectal cancer screening programme: results from the first 2.6 million invitations in England, Int. J. Epidemiol. 40 (2011) 712–8. doi:10.1093/ije/dyr008.

[26] M. Burnell, A. Gentry-Maharaj, A. Ryan, S. Apostolidou, M. Habib, J. Kalsi, S. Skates, M. Parmar, M.W. Seif, N.N. Amso, K. Godfrey, D. Oram, J. Herod, K. Williamson, H. Jenkins, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, S. Leeson, D. Cruickshank, S. Campbell, L. Fallowfield, I. Jacobs, U. Menon, Impact on mortality and cancer incidence rates of using random invitation from population registers for recruitment to trials, Trials. 12 (2011) 61. doi:10.1186/1745-6215-12-61.

[27] M. Arnold, M.S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global patterns and trends in colorectal cancer incidence and mortality, Gut. 66 (2017) 683–691. doi:10.1136/gutjnl-2015-310912.

[28] A. Gentry-Maharaj, E.-O. Fourkala, M. Burnell, A. Ryan, S. Apostolidou, M. Habib, A. Sharma, M. Parmar, I. Jacobs, U. Menon, Concordance of National Cancer Registration with self-reported breast, bowel and lung cancer in England and Wales: a prospective cohort study within the UK Collaborative Trial of Ovarian Cancer Screening, Br. J. Cancer. 109 (2013) 2875–2879. doi:10.1038/bjc.2013.626.

[29] J.C. Thorn, E. Turner, L. Hounsome, E. Walsh, J.L. Donovan, J. Verne, D.E. Neal, F.C. Hamdy, R.M. Martin, S.M. Noble, Validation of the Hospital Episode Statistics Outpatient Dataset in England, Pharmacoeconomics. 34 (2016) 161–8. doi:10.1007/s40273-015-0326-3.

[30] M.M. Bergmann, E.E. Calle, C.A. Mervis, H.L. Miracle-McMahill, M.J. Thun, C.W. Heath, Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries, Am. J. Epidemiol. 147 (1998) 556–62.

[31] A. Parikh-Patel, M. Allen, W.E. Wright, California Teachers Study Steering Committee, Validation of self-reported cancers in the California Teachers Study, Am. J. Epidemiol. 157 (2003) 539–45.

[32] R.J. Mitchell, D. Brewster, H. Campbell, M.E.M. Porteous, A.H. Wyllie, C.C. Bird, M.G. Dunlop, Accuracy of reporting of family history of colorectal cancer, Gut. 53 (2004) 291–5. doi:10.1136/gut.2003.027896.

[33] S.H. Lo, J. Waller, C. Vrinten, J. Wardle, C. von Wagner, Self-Reported And Objectively Recorded Colorectal Cancer Screening Participation In England, J. Med. Screen. 23 (2016) 17–23. doi:10.1177/0969141315599015.

[34] Parliamentary Office of Science & Technology, Electronic Health Records, 2016.

Chapter 4. Development of A Prediction Tool for Informing Risk-Stratified Colorectal Cancer Screening for UK Women

A Retrospective Case-Control Study Nested Within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

4.1 Background

Screening programmes are viable if their benefits outweigh the harms**.** On the benefits, asymptomatic CRCs detected before evolving invasive and metastatic traits are highly treatable and have improved survival outcomes [1]. Benefits withstanding, a small minority of CRC screens cause harm from the anxiety caused by false-positive screening results or from complications of diagnostic colonoscopy [2,3]. Risk-prediction tools may help to redress the balance NHS cancer screening programs in favour of benefits [4].

Current screening for CRC partitions susceptibility solely on age; assuming those aged 60–74 (to be lowered to 50–74 in alignment with Scotland [5]) would benefit equally from fixed, biennial screening with a standardised screening test. Age, however, cannot account for all variation in risk, and several studies have demonstrated how factoring in genetic and non-genetic variables in addition to age can enhance the efficiency of screening [6,7] and diagnostic referrals [8].

In the current era where genotyping is not yet routine throughout healthcare, tools exploiting non-genetic risks that are recorded within routine EHRs would surpass many operational challenges [9]. Moreover, a study of identical and non-identical twins has demonstrated that non-genetic factors account for two-thirds of CRC susceptibility [10]. The present population-based, retrospective case-control study, therefore, aimed to derive a non-genetic risk score (NGRS) that could inform risk-stratified screening, using variables routinely recorded within EHRs. The NGRS was evaluated on its ability to reclassify the risk of a bootstrap sampled in comparison to the current eligibility on age alone.

4.2 Methods

Study design. The present population-based, retrospective cohort study was nested within the UKCTOCS. The study aligns with the TRIPOD statement for reporting multivariable risk-prediction models [11].

Study cohort. Subjects were 202,638 post-menopausal women aged 50–74 recruited from NHS patient registers to one of thirteen regional UKCTOCS centres in England, Northern Ireland, & Wales during 2001–05 [12]. Women with previous bilateral oophorectomy or ovarian malignancy, increased risk of familial ovarian cancer, or an active non-ovarian malignancy were excluded from recruitment to the UKCTOCS. One half of the UKCTOCS cohort were assigned as non-interventional controls, a quarter to received ultrasound screening, and a quarter receiving screening by blood test and ultrasound combined. UKCTOCS subjects were eligible for this study if they consented to their involvement in secondary studies. Subjects who reported residence other than in England, Northern Ireland, or Wales were excluded.

Data sources

Electronic Health Records. Subjects were linked to their EHRs (cancer & death registries, secondary-care records, and the National Cancer Intelligence Network (NCIN) registry) by their NHS number. Secondary-care records were available for England (HES) and Wales (Patient Episode Database for Wales) residents only. NCIN data were available for women in England. The latest update received for each EHR is presented in Table 4–1.

Table 4–1: Coverage of Electronic Health Records.

Electronic Health Record	Latest update
CR (England & Wales)	6 February 2018
CR (Northern Ireland)	15 November 2017
DC (England & Wales)	6 February 2018
DC (Northern Ireland)	30 August 2017
HES (England)	8 January 2018
PEDW (Wales)	11 December 2017

Abbreviations: CR, cancer registration; DC, death certificate; HES, Hospital Episode Statistics; PEDW, Patient Episode Database for Wales.

Self-reported questionnaires. Subjects anthropometrics, medical history, and lifestyle data were self-reported I) in-person via a recruitment questionnaire (RQ; (Supplementary Figure S1) completed at enrolment during 2001–05 or via postal return of the II) follow-up questionnaire I (FUQ I; (Supplementary Figure S2)) sent to all women at 3–5-years post-randomisation and/or III) follow-up questionnaire II (FUQ II; (Supplementary Figure S3)) sent to all surviving women in April 2014. The RQ, FUQ I, and FUQ II was completed by 100% (202,232/202,232), 71% (144,216/202,232), and 41%

(83,398/202,232) of consenting women residing in England, Wales or Northern Ireland, respectively. Postal addresses were updated throughout the trial.

Outcomes. It was discerned in the previous chapter that cancer & death registries & HES used in adjunct can identify 98% of all events with a PPV of 92%. Cases were therefore defined as women who had a positive notification of CRC (ICD9 153–4 or ICD10 C18–20) diagnosed since birth as per these records. Few self-reported but verified cancers ascertained in Chapter 3 were also included. Women without notification from these records were defined as controls. Age at right censorship was calculated as the years between date of birth (self-reported at recruitment) and date of cancer diagnosis (in weight order of clinical confirmation (date of diagnosis), cancer registration (date of diagnosis), secondary-care records (date of first episode), or death registration (date of death)) for cases or date of latest EHR update (or date of death if before) for controls. The latest EHR updates were 6 February 2018 for England, 15 November 2017 for Northern Ireland, and 6 February 2018 for Wales (Table 4–1).

Covariates. Covariates were discerned from EHRs (Benign) or were self-reported via questionnaire(s). Benign was diagnoses of carcinoma *in situ* (ICD D010–12) or benign adenomatous polyp (ICD D12). Diagnoses of carcinoma *in situ* were identified from CR, secondary-care datasets (HES for England, PEDW for Wales), and NCIN (England only). Diagnoses of benign adenomatous polyp were identified from CR and secondary-care datasets (HES/PEDW). All other covariates were self-reported in response to questionnaire items transcribed verbatim in Table 4–2.

Preselection of covariates. Modelled covariates were limited to those I) routinely recorded in EHRs and II) that have robust evidence from meta-analyses or randomised controlled trials for their inclusion. Thirteen covariates were excluded for either not being recorded by EHRs (Dress size, Education, Family history of cancer, Marital Status), having no evidence supporting their inclusion (Infertility treatment [13], Menarche [14], Parity [15], Tamoxifen [16], Tubal ligation [17] or no relevant studies conducted (Cancer history (other than colorectal), Osteoarthritis, Stroke), or for not capturing adequate information (Thyroid disease—hyperthyroidism and use of thyroid replacement hormones increase the risk of CRC while untreated hypothyroidism decreases risk [18]). Twenty-six covariates were preselected for modelling. Evidence for their inclusion,

source, and entry or engineering are presented in Table 4–2. These are broadly classified as:

—*Anthropometrics* (BMI, Ethnicity, Height, Hip/Waist circumferences, Waist-to-hip ratio) that are generally taken by a practice nurse (or self-identified for ethnicity) on registry to a primary care practice. Recency of BMI and Hip/Waist circumference measurements thus has obvious implications for risk prediction.

—*Lifestyle choices* (Alcohol, Smoking) whose collection is encouraged across primary and secondary care and is primarily used for healthcare promotion.

—*Events & morbidities* (Benign, Arthritis, Cholesterol, COPD, Diabetes, Gravidity, Heart disease, Hypertension, IBD, Kidney disease, Liver disease, Osteoporosis) encoded as ICD codes throughout primary and secondary care.

—*Treatments & surgery* (Aspirin, HRT, Hysterectomy, OCP use/years, Statin) encoded as SNOWMED-CT codes throughout primary and secondary care.

Data analyses.

Incidence rate. The incidence rate was calculated as: (Σ cases/(Σ person-years follow-up/100,000)). The incidence rate was then compared as a percentage of the Σ of cases given the European age-standardised rate for females being 57.2 incidences per 100,000 person-years [19]. Use of an age-standardised rate allows for comparisons irregardless of the age structures.

Entry & engineering of covariates. Entry and engineering of all preselected covariates are summarised in Table 4–2. Briefly, BMI, Gravidity (pregnancies < 6 months), Hip circumference, Height, OCP use (years), Waist circumference, and Waist-to-hip ratio were coded as continuous variables. All other covariates were categorical. Missing continuous data were entered as missing (excluding pregnancy data, which was coded as 0). Categorical data were assumed absent if the questionnaire item(s) was/were unfilled (excluding Alcohol & Smoking, which were entered as missing if unfilled) and missing if questionnaire unreturned. Outlying continuous data were treated as missing if considered implausible. Outliers were defined as data ± three standard deviations of the mean (~ 99.7% of a Gaussian population) calculated for the UKCTOCS cohort or from the 2015 Health Survey England [20], whichever was more appropriate. Pregnancies were taken at face value. BMI and Waist-to-hip ratio were calculated after removal of outlying data for Weight & Height and Hip circumference & Waist circumference, respectively.

Table 4–2: Twenty-six engineered features.

Feature	Current evidence	Criteria/Question phrasing	Data entry
Electronic Health Records			
Benign	Meta-analysis [21]: polypectomy patients have increased risk of subsequent CRC (RR 2.52 (95% CI 1.35–4.73)).	ICD code for carcinoma *in situ* (D010–12) or benign adenomatous polyps (ICD D12) in secondary-care and National Cancer Intelligence Network datasets.	Binary. Coded 1 if previous diagnosis of benign adenomatous polyp or carcinoma *in situ* recorded in EHR; 0 if otherwise.
Recruitment questionnaire completed by 100% of women (202,232/202,232)			
BMI	Meta-analysis [22]: obesity associated with increased risk (RR 1.33 (95% CI 1.25–1.42)).	'Your height (cm) or (In)' and 'Your weight (kg) or (lb).'	Continuous (Kg m^{-2}). Imperial (In or lb) measurements converted to metric (cm/kg). BMI (Weight/((Height/100)^2)) calculated from height (m) and weight (Kg) self-reported at recruitment to the UKCTOCS (after removal of Weight/Height outliers).
Ethnicity	National consensus in England [23]: Relative to White (British/Irish/Other), incidence of CRC lower in Black (Rate ratio: 0.70 (95% CI 0.66–0.74)), Chinese (0.68 (95% CI 0.59–0.78)), and Asian subgroups (0.45 (0.43–0.47)).	'Ethnic group, please place an "X" in the appropriate box. (If you are descended from more than one ethnic or racial group, please select the group you consider you belong to or chose "Any other ethnic origin").'	Nominal: categorised as White, Black (African or Caribbean or Black-other), Asian (Indian or Pakistani or Bangladeshi), Chinese or Other.
Gravidity	No meta-analyses, but some evidence [24] for a decreased risk due to gravidity ≥ 3 (OR 0.64 (95% CI 0.43–0.94)).	'How many pregnancies have you had which ended before they reached 6 months (including miscarriages, ectopic pregnancies)?'	Continuous (n). Nulligravida coded as 0. Missing data coded as 0.

Height	Prospective cohort study [25]: risk of colon (RR 1.25 (95% CI 1.19–1.30) and rectal (RR 1.14 (95% CI 1.07–1.22)) cancer increased per 10 cm increase in height.	Your height (cm) or (In).'	Continuous (cm). Imperial In converted to metric cm if necessary.
OCP (ever use)	Meta-analysis [26]: ever OCP use associated with decreased risk (RR 0.81 (95% CI 0.72–0.92)).	'Have you ever taken the oral contraceptive pill?'	Binary: 1 if self-reported or OCP duration (below) > 0) and 0 if item unfilled. UKCTOCS stored infilled item as 0.
OCP (duration)	Meta-analysis [26]: long-term use ≥5 years associated with decreased risk (RR 0.86 (95% CI 0.74–1.00)), while no advantage observed with <5-years use (RR 0.88 (95% CI 0.77–1.01)).	'…how many years in total did you take the pill?'	Continuous (years). Never use /uncomplete item coded as 0.
Follow-up questionnaire I completed by 71% of women (144,216/202,232)			
Alcohol	Meta-analysis [27]: dose-dependent increase in risk (1.52 RR (95% CI 1.27–1.81)).	'Approximately how much alcohol on average do you drink each week? (One drink = a glass of wine, half a pint of lager or cider, a measure of spirits).'	Ordinal: 0/< 1/≥ 1 drink per day. Note that, while easier to self-report, these alcohol measurements are not necessarily standardised and do not equate to 1 alcohol unit.
Smoking	Meta-analysis [28]: current and former smokers have increased risk (RR 1.20 (95% CI 1.10–1.30) and RR 1.18 (95% CI 1.12–1.25)), respectively.	'Have you ever been a smoker? If you answered yes to the above, please answer the following questions: How many years in total have you smoked for? During these years, how many cigarettes on average did you smoke per day?	Ordinal: Ever tobacco users defined as that self-reporting either: ever use, ≥ 1 year of use, or ≥ 1 cigarette a day. Never were those who did not satisfy the above criteria (if reported).

Follow-up questionnaire II completed by 41% of women (83,398/202,232)			
COPD	No meta-analyses; two case-control studies [29,30] show that COPD associated increased risk of adenomatous polyps (OR 2.06 (95% CI 1.13–3.76) but not CRC (OR 1.14 (95% CI 0.73–1.77)).	'Do you have/are being treated for any of the following conditions? Chronic Obstructive Pulmonary Disease.'	Binary: 1 if self-reported, 0 if item unfilled, and missing if questionnaire was unreturned.
Hip circumference	Meta-analysis [31]: each 8-cm (~3 In) increase in hip circumference increased risk by 15% (HR 1.15 (95% CI 1.01–1.32)).	'Please measure and enter the following measurements: Your hips (inches).'	Continuous (inches). Missing if questionnaire unreturned or item was unfilled.
IBD	Meta-analysis [32]: IBD patients—Ulcerative & Crohn's colitis—have 1.7-fold increased SIR (95% CI 1.2-2.2).	Do you have/are being treated for any of the following conditions? Inflammatory bowel disease.'	Binary: 1 if self-reported, 0 if item unfilled, and missing if questionnaire was unreturned.
Kidney disease	Meta-analysis [33]: chronic kidney diseases increases risk (SIR 1.18 (95% CI 1.01–1.37)).	'Do you have/are being treated for any of the following conditions? Kidney disease.'	Binary: 1 if self-reported, 0 if item unfilled, and missing if questionnaire was unreturned.
Liver disease	Meta-analysis [34]: patients with chronic liver diseases—hepatitis & cirrhosis—have an increased risk (SIR 2.06 (95% CI 1.46–2.90)). Primary sclerosing cholangitis markedly increases risk (SIR 6.70 (95% CI 3.48–12.91)).	'Do you have/are being treated for any of the following conditions? Liver disease.'	Binary: 1 if self-reported, 0 if item unfilled, and missing if questionnaire was unreturned.
Aspirin [low-dose)	Meta-analysis [35]: long-term aspirin users have lower risk (HR 0.76 (95% CI 0.63–0.94)).	'Have you ever taken any of the following medications? Low-dose aspirin.'	Binary: 1 if self-reported, 0 if item unfilled/stated no, and missing if questionnaire was unreturned.

Statin	Meta-analysis [36]: statin use associated with decreased risk (RR 0.91 (95% CI 0.87–0.96)).	'Have you ever taken any of the following medications? Statins.'	Binary: 1 if self-reported, 0 if item unfilled/stated no, and missing if questionnaire was unreturned.
Waist circumference	Meta-analysis [22]: those in the highest category of waist circumference had increased risk (RR 1.46 (95% CI 1.33–1.60)) to those in the lowest category (categories study-specific).	'Please measure and enter the following measurements: Your waist (inches).'	Continuous (inches). Missing if questionnaire unreturned or item was unfilled.
Waist-hip ratio	Meta-analysis [37]: those in the highest category of waist-hip ratio had increased risk (RR 1.39 (95% CI 1.25–1.53)) to those in the lowest category (categories study-specific).	See Hip circumference & Waist circumference.	Continuous (ratio): waist circumference divided by hip circumference.
Recruitment questionnaire & Follow-up questionnaire I			
HRT (ever use)	Meta-analysis [38]: RRs associated with current use of oestrogen and oestrogen-progesterone were 0.70 (95% CI 0.57–0.85) and 0.80 (95% CI 0.69–0.93), respectively.	RQ: 'Are you currently on Hormone Replacement Therapy (HRT)?' FUQ I: 'Are you currently taking HRT?'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported.
Recruitment questionnaire & Follow-up questionnaires I & II			
Hysterectomy	Meta-analysis [39]: hysterectomy associated with increased risk (RR 1.24 (95% CI 1.17–1.32)).	RQ: 'Have you ever had a hysterectomy (removal of the womb)?' FUQ I: 'Since joining UKCTOCS have you had any of the following operations? Hysterectomy/Removal of womb.' FUQ II: 'Have you had any of the following procedures since joining	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported.

		UKCTOCS? Operation to remove your womb (hysterectomy).'	
Follow-up questionnaires I & II (75% (152,041/202,232) completed at least one)			
Arthritis (rheumatoid)	Meta-analysis [40]: rheumatoid arthritis patients have decreased risk (SIR 0.78 (95% CI 0.71–0.86)).	FUQ I: 'Do you have/are you being treated for any of the following conditions? Rheumatoid arthritis.' FUQ II: 'Do you have/are being treated for any of the following conditions? Rheumatoid arthritis.'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.
Diabetes	Meta-analysis [41]: diabetes associated with increased risk (RR 1.22 (95% CI 1.19–1.26)).	FUQ I: 'Do you have/are you being treated for any of the following conditions? Diabetes.' FUQ II: 'Do you have/are being treated for any of the following conditions? Diabetes.'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.
Heart disease	Meta-analysis [42]: Patients with [ischemic] heart disease had increased risk (OR 1.87 (95% CI 1.38–2.54)).	FUQ I: 'Do you have/are you being treated for any of the following conditions? Heart disease, e.g. heart attack, angina.' FUQ II: 'Do you have/are being treated for any of the following conditions? Heart disease (e.g. heart attack, angina).'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.
Cholesterol (high-blood)	Meta-analysis [43]: high total cholesterol (RR 1.11 (95% CI 1.01–1.21) associated with increased risk. RRs associated with high and low-density lipoprotein cholesterol were	FUQ I: 'Do you have/are you being treated for any of the following conditions? High blood cholesterol.' FUQ II: 'Do you have/are being treated for any of the following conditions? High blood cholesterol.'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.

	0.84 (95% CI 0.69–1.02) and 1.04 (95% CI 0.60–1.81), respectively.		
Hypertension	Meta-analysis [44]: hypertension associated with increased risk (RR 1.09 (95% CI 1.01–1.18)).	FUQ I: 'Do you have/are you being treated for any of the following conditions? High blood pressure.' FUQ II: 'Do you have/are being treated for any of the following conditions? High blood pressure.'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.
Osteoporosis	No meta-analyses, but bone mineral density inversely proportional to colon cancer risk [45], and there's some evidence [46] and osteoporosis increases the risk of colorectal adenoma (OR 1.59 (95% CI 1.00–2.52)).	FUQ I: 'Do you have/are being treated for any of the following conditions? Osteoporosis.' FUQ II: 'Do you have/are being treated for any of the following conditions? Osteoporosis.'	Binary: 1 if self-reported in either questionnaire, 0 if not self-reported, and missing if both questionnaires were unreturned.

Cox proportional hazards regression. Missing data were imputed 5, 10, and 20 times using predictive mean matching (continuous), logistic regression (binary categorical), and Bayesian polytomous regression (> 2 categories) incorporating all preselected covariates and outcome as predictors. Final analyses were based on the least number of imputations from which stable standard errors of the coefficients were obtained as per the Kruskal-Wallis test. Covariates with ≥ 50% missingness was excluded from analysis. Models were fitted using Cox proportional hazards regression fitted to each imputed dataset and were aggregated as per Ruben's rules [47]. We used Efron's approximation events for reasons of computing efficiency given the dataset had few tied events [48]. The time-to-event outcome was age at date of diagnosis or right censorship for cases and controls, respectively. All features preselected *a priori* were retained in the final model given the robust evidence for their inclusion. Assumptions of proportional hazards were verified on the correlation between transformed Age (time variable) and scaled Schoenfeld residuals. Heart disease had disproportional hazards over time and we therefore stratified as per Heart disease status. Linearity of continuous covariates was verified by plotting covariates individually against Martingale residuals.

Model validation. The model was validated on a bootstrap resample of 138,900 subjects (2,678 CRCs) with complete data in all preselected covariates.

Non-genetic risk score. A NGRS was derived from the linear predictor form of the Cox proportional hazards. Cox proportional hazards defines the hazard function at time t for a given set of covariates x ($h(t|x)$) and is calculated given the baseline hazard function ($h0(t)$)—the time-related component—and exponentiated linear predictor ($e^{\beta 1X1+\beta 2X2+\beta nXn}$)—the covariate-related component—as:

$$h\,(t|xi) = h0\,(t)\,e^{\beta 1Xi+\beta 2Xi+\beta nXi} \quad \textbf{(1)}$$

With the computed regression coefficients (β) and covariates (X), the linear predictor was calculated as:

$$\beta 1Xi + \beta 2Xi + \beta nXi \quad \textbf{(2)}$$

The NGRSs were calculated from the addition of β coefficients multiplied by covariates.

Risk reclassification. The proportion of potentially detectable CRCs (CRCs eligible for screening/total CRCs) and the number of unnecessary screens (controls eligible for screening/total number of controls) was calculated for the subset of women with complete age data assuming screening eligibility on the revised criteria (50–74-years). We then compared the same proportions after classifying the risk of the bootstrap resample using the NGRS with the linear predictor threshold equivalent to the sensitivity of age-only eligibility (i.e. to minimise harms). The risk gradient of the NGRS was calculated as the exponentiated β coefficient for the 80th percentile minus the exponentiated β coefficient for the 20th percentile.

Statistical analyses. All analyses were made with R version 3.3.2 [49] running epiR [50], ggplot2 [51], mice [52], miceadds [53], rms [54], and survival [55] packages. A p value < 0.05 was considered statistically significant. Normal distributions were assumed for all continuous variables as per the central limit tendency of large datasets. The online InteractiVenn tool was used to produce the Venn diagram [56].

4.3 Results

Outcomes. Two-hundred & seventy-three women who wished not to be involved in secondary studies and 133 women who resided outside England, Northern Ireland or Wales were excluded. Of the remaining 202,232 women, 4,134 (2.0%) were defined as cases, and the remaining 198,098 (98.0%) were controls with no evidence of any diagnoses. Subjects who developed cancer were 68.7 years on average (range 23–89 years). EHRs informed 4,127 of 4,134 cancers, while the remaining seven were self-reported (but verified in Chapter 3). Of these 4,127 cancers, 1,467 (36%) were informed by one EHR (909 cancer registries, 492 HES, 66 death registries), 1,962 (48%) by two EHRs, and 698 (17%) by all three EHRs (Figure 4–1). Date of diagnosis for the 4,134 events was informed by clinician's confirmation (n 533), cancer registry (n 2,971), secondary-care records (n 564) or death registry (n 66). A total of 20,833 controls exited the study early due to death. The incidence rate given 4,134 cancers after 15,109,459 person-years follow-up was 27.4 per 100,000 person-years (4,134/(1,5109,459/100,000)). Accordingly, 48% (27.4/57.2) of the expected incidences were observed.

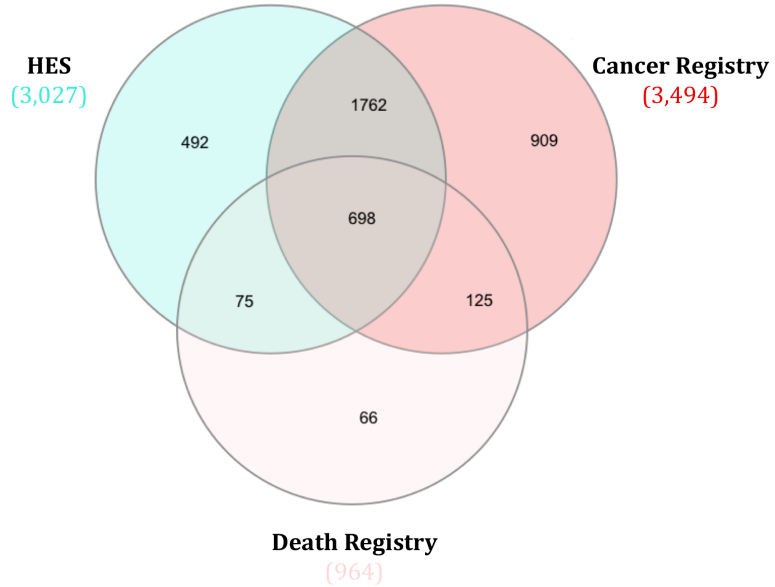


Figure 4–1: Electronic Health Records ascertained 4,127 of 4,134 cancers. Five-hundred & fifteen cancers were verified by the treating clinician in Chapter 3. Seven unaccounted for cancers were self-reported but verified.

The distribution of Age at diagnosis or right censorship is presented in Figure 4–2. Data were missing for one woman who did not provide their date of birth. The mean age at diagnosis of CRC was 69 and ranged from 23–88 years. The mean age at right censorship for controls was 74 and ranged from 50–91 years.

Figure 4–2: Distribution of ages at diagnosis of CRC (cases) or right censorship (controls).

Covariates. Covariates were self-reported by subjects in-person or via postal questionnaire. All 202,232 women completed the recruitment questionnaire in-person, while 71% (144,216/202,232) and 41% (83,398/202,232) completed by post their FUQ I and FUQ II, respectively. Thirty-seven percent (75,573/202,232) of women returned both FUQ I & II, 38% (76,468/202,232) women either FUQ I or II, while 25% (50,191/202,232) women returned neither. There were 1,497 outliers for BMI (< 5.1 or > 49.1 kg m^{-2} (Health Survey England)), 718 for height (< 80.7 or > 226.6 cm (Health Survey England)), 794 for Hip circumference (< 26.7 or > 53.4 inches (UKCTOCS)), 529 for Waist circumference (< 16.7 or > 49.0 inches (UKCTOCS)), and 519 for Wait-to-hip ratio (< 0.59 or > 1.05 (UKCTCOS)). These were processed as being missing. A summary of descriptive and frequency statistics for covariates are summarised in Table 4–3.

Table 4–3: Summary characteristics.

Covariate	All (n 202,232)	Controls (n 198,098)	Cancers (n 4,134)
Alcohol			
≥ 1 drink per day	33,241 (16%)	32,650 (16%)	591 (14%)
< 1 drink per day	75,305 (37%)	73,874 (37%)	1,431 (35%)
Sober	33,985 (17%)	33,287 (17%)	698 (17%)
Missing data	59,701 (30%)	58,287 (29%)	1,414 (34%)
Arthritis (Rheumatoid)			
Yes	9563 (5%)	9,383 (5%)	180 (4%)
No	142,478 (70%)	139,767 (71%)	2,711 (66%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)
Aspirin (low-dose)			
Yes	15,082 (7%)	14,815 (7%)	267 (6%)
No	68,316 (34%)	67,324 (34%)	992 (24%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)
Benign			
Yes	9,082 (4%)	8,106 (4%)	976 (24%)
No	193,150 (96%)	189,992 (96%)	3,158 (76%)
Missing data	0 (0%)	0 (0%)	0 (0%)
BMI (Kg m^{-2})			
Mean	26.6	26.6	26.8
SD	4.8	4.7	4.9
Missing data	2,507 (1%)	2,455 (1%)	52 (1%)
Cholesterol			
Yes	46,348 (23%)	45,484 (23%)	864 (21%)
No	105,693 (52%)	103,666 (52%)	2,027 (49%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)

COPD			
Yes	2,576 (1%)	2,525 (1%)	51 (1%)
No	80,822 (40%)	79,614 (40%)	1,208 (29%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)
Diabetes			
Yes	10,707 (5%)	10,457 (5%)	250 (6%)
No	141,334 (70%)	138,693 (70%)	2,641 (64%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)
Ethnicity			
Asian	1,473 (1%)	1,455 (1%)	18 (0%)
Black	2,748 (1%)	2,705 (1%)	43 (1%)
Chinese	373 (0%)	364 (1%)	9 (0%)
Other	1,685 (1%)	1,667 (%)	18 (0%)
White	194,912 (96%)	190,892 (96%)	4,020 (97%)
Missing data	1,041 (1%)	1,015 (1%)	26 (1%)
Gravidity (n)			
Mean	0.5	0.5	0.4
SD	0.9	0.9	0.8
Missing data	0 (0%)	0 (0%)	0 (0%)
Heart disease			
Yes	12,045 (6%)	11,804 (6%)	241 (6%)
No	139,996 (69%)	137,346 (69%)	2,650 (64%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)
Height (cm)			
Mean	161.6	161.6	162.0
SD	6.8	6.8	7.0
Missing data	1,218 (1%)	1,188 (1%)	30 (1%)
Hip circumference (In)			
Mean	40.0	40.0	40.1
SD	3.5	3.5	3.5
Missing data	131,710 (65%)	128,600 (65%)	3110 (75%)
HRT			
Yes	39,835 (20%)	39,157 (20%)	678 (16%)
No	162,397 (80%)	158,941 (80%)	3,456 (84%)
Missing data	0 (0%)	0 (0%)	0 (0%)
Hypertension			
Yes	56,655 (28%)	55,491 (28%)	1,164 (28%)
No	95,386 (47%)	93,659 (47%)	1,727 (42%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)
Hysterectomy			
Yes	42,163 (21%)	41,325 (21%)	838 (20%)
No	160,069 (79%)	156,773 (79%)	3,296 (80%)
Missing data	0 (0%)	0 (0%)	0 (0%)
IBD			
Yes	2,923 (1%)	2,878 (1%)	45 (1%)
No	80,475 (40%)	79,261 (40%)	1,214 (29%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)

Kidney disease			
Yes	1,476 (1%)	1,447 (1%)	29 (1%)
No	81,922 (41%)	80,692 (41%)	1,230 (30%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)
Liver disease			
Yes	889 (0%)	871 (0%)	18 (0%)
No	82,509 (41%)	81,268 (41%)	1,241 (30%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)
OCP use			
Yes	120,483(60%)	118,375 (60%)	2,108 (51%)
No	81,749 (40%)	79,723 (40%)	2,026 (49%)
Missing data	0 (0%)	0 (0%)	0 (0%)
OCP use (years)			
Mean	4.1	4.1	3.5
SD	5.7	5.7	5.6
Missing data	1,810 (1%)	1,778 (1%)	32 (1%)
Osteoporosis			
Yes	15,864 (8%)	15,542 (8%)	322 (8%)
No	136,177 (67%)	133,608 (67%)	2,569 (62%)
Missing data	50,191 (25%)	48,948 (25%)	1,243 (30%)
Smoking			
Ever	62,823 (31%)	61,565 (31%)	1,258 (30%)
Never	81,147 (40%)	79,639 (40%)	1,508 (36%)
Missing data	58,262 (29%)	56,894 (29%)	1,368 (33%)
Statin			
Yes	27,338 (14%)	26,898 (14%)	440 (11%)
No	56,060 (28%)	55,241 (28%)	819 (20%)
Missing data	118,834 (59%)	115,959 (59%)	2,875 (70%)
Waist circumference (In)			
Mean	32.8	32.8	32.8
SD	4.5	4.5	4.5
Missing data	131,218 (65%)	128,118 (65%)	3100 (75%)
Waist-to-hip ratio			
Mean	0.82	0.82	0.82
SD	0.07	0.07	0.07
Missing data	133,046 (66%)	129,915 (66%)	3,131 (76%)

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; HRT, hormone replacement therapy; IBD, inflammatory bowel disease; OCP, oral contraceptive pill.

Cox proportional hazards regression. Missingness by variable ranged from 0–66% (Figure 4–3). The nine variables self-reported in the FUQ II had the greatest missingness (59–66%). These variables were therefore excluded given that the reliability of their imputation could not be assured. After removal of the nine features with ≥ 59% missingness, 138,900 of 202,232 (69%) subjects had complete data. Controls had a greater

proportion of complete data (69% (136,241/190,098)) than cases (64% (2,659/4,134)). The number of imputations required to overcome the uncertainties with imputation was then determined based on errors of feature coefficients. The sum of pooled standard errors obtained from 5, 10, and 20 imputations were 3.22, 3.28, 3.17, respectively, and was not statistically different from one another (p 0.972). Accordingly, five imputations were used for all subsequent analyses.

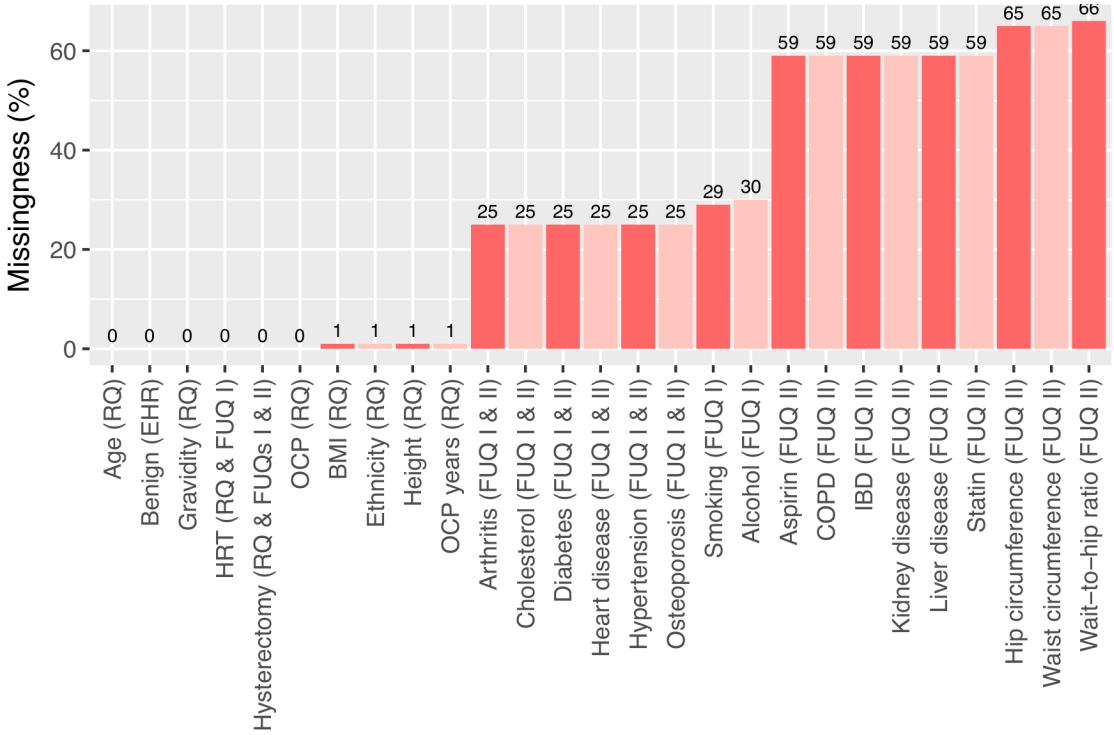


Figure 4–3: Missingness in variables for 202,232 women. All women completed the RQ, while 71% (144,216/202,232) and 41% (83,398/202,232) completed their FUQ I and II, respectively; 37% (75,573/202,232) returned both; 38% (76,468/202,232) returned at least one, and 25% (50,191/202,232) returned neither). Abbreviations: RQ, recruitment questionnaire; FUQ, follow-up questionnaire.

The initial Cox Proportional Hazards model violated the assumption of proportional hazards (p 0.002) owing to disproportionable hazards over time for Cholesterol (p 0.004), Heart disease (p 5e-4), and Osteoporosis (p 0.011). The adjusted hazard ratios associated with the NGRS is presented in Table 4–4. The assumption of proportional hazards was met by stratifying according to Heart disease status (0.186), which explains why Heart disease does not have a HR reported.

Table 4–4 Adjusted Hazard Ratios.		
Covariate	**Adjusted HR**	**95% CI**
Alcohol		
Sober	1.00	—
< 1 drink per day	0.98	0.90–1.07
≥ 1 drink per day	0.97	0.84–1.13
Arthritis		
No	1.00	—
Yes	0.87	0.72–1.06
Benign		
No	1.00	—
Yes	6.47	6.02–6.96
BMI (Kg m^{-2})		
per unit increase	1.01	1.00–1.02
Cholesterol		
No	1.00	—
Yes	0.84	0.77–0.91
Diabetes		
No	1.00	—
Yes	1.14	0.98–1.31
Ethnicity		
White	1.00	—
Asian	0.85	0.53–1.36
Black	0.87	0.65–1.19
Chinese	1.74	0.90–3.35
Other	0.70	0.44–1.11
Gravidity (n)		
per unit increase	0.98	0.95–1.02
Height (cm)		
per unit increase	1.01	1.01–1.02
HRT		
No	1.00	—
Yes	1.04	0.95–1.13
Hypertension		
No	1.00	—
Yes	0.94	0.86–1.01
Hysterectomy		
No	1.00	—
Yes	0.91	0.84–0.99
OCP		
No	1.00	—
Yes	1.00	0.92–1.08
OCP use (years)		
per unit increase	1.00	1.00–1.01
Osteoporosis		
No	1.00	—
Yes	0.95	0.84–1.06

Smoking		
Never	1.00	—
Ever	1.10	1.02–1.18

Abbreviations: BMI, Body Mass Index; HRT, Hormone Replacement Therapy; OCP, Oral Contraceptive Pill.

Risk reclassification. We then sought to reclassify the risk of a bootstrap resample of 138,900 women who had complete data in all preselected covariates. The incidence rate within the bootstrap resample given 2,678 cases in 10,400,166 person-years follow-up was 25.7 cases per 100,000 person-years (2,678/(10,400,166/100,000)). Under the current screening eligibility based on age only (50–74 years), 65.6% (2,711/4,134) of CRCs would be potentially detectable while 48.6% (96,323/198,097) of screens will be unnecessary. To potentially detect an equal proportion of cancers (65.6% (1,757/2,678)), 5.6% more unnecessary screens (54.2% (73,888/136,222)) would be undertaken if eligibility was determined with a NGRS. The age of those eligible were 33–87-years.

Table 4–5: Reclassification of screening eligibility based on a) age-only and b–c) non-genetic risk scores.

a

50–74	**Outcome +**	**Outcome –**
+	2,711 TP	96,323 FP
–	1,423 FN	101,774 TN

b

HR > 10.32	**Outcome +**	**Outcome –**
+	1,757 TP	73,888 FP
–	921 FN	62,334 TN

a) classification of UKCTOCS women with complete data in all preselected features (n 202,231 (1 subject with missing age excluded)). b) classification of bootstrap resample of subjects with complete data in all preselected covariates (n 138,900) using a NGRS.

The distribution of risk associated with the NGRS is presented in Figure 4–4. Relative to the geometric mean (50th percentile), those in the 99th and 90th percentiles had a 64-fold (HR 64.44 (99 percentile 74.96 – 50 percentile 10.52)) and 3-fold (HR 2.57 (90 percentile 13.09 – 50 percentile 10.52)) greater hazards of developing CRC. Women in the top quintile of risk had three-fold greater hazards of developing CRC to women in the bottom quintile (HR 2.98 (Q5 12.04 – Q1 9.06)).

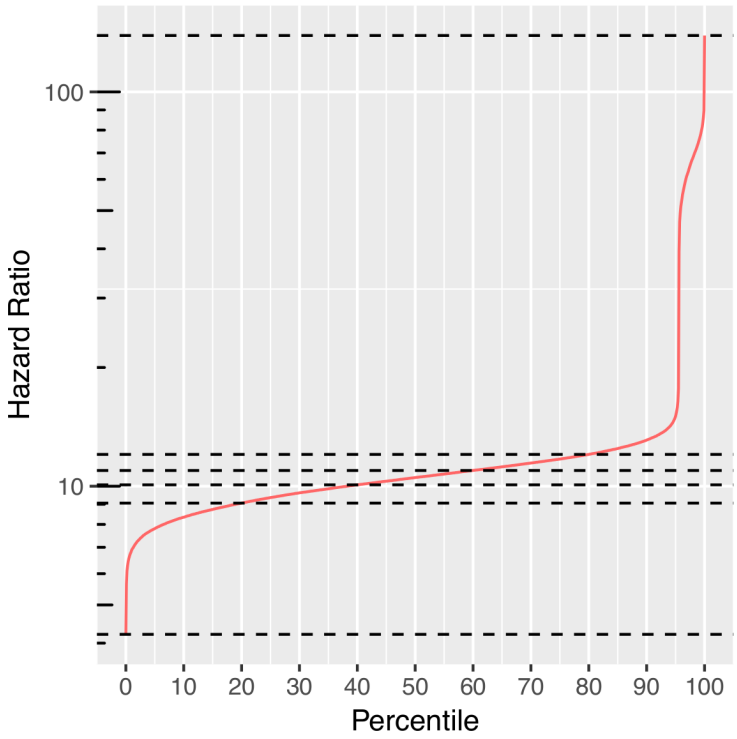


Figure 4–4: Risk-gradient of NGRS applied to a bootstrap resample (n 138,900). Quintile bands are demarcated with dashed lines. The HRs for major percentiles were as follows: 1%, 6.88; 5%, 7.81; 10%, 8.34; 20%, 9.06; 30%, 9.62; 40%, 10.08; 50%, 10.52; 60%, 10.96; 70%, 11.45; 80%, 12.04; 90%, 13.09; 95%, 15.05; 99%, 74.96; 100%, 138.98.

4.4 Discussion

Principle findings. The UK National Screening Committee has recently recommended that screening should be provided from the age of fifty—10-years earlier than currently provided for [5]. Despite occult blood testing being highly specific [57], the harms of screening and the burden to endoscopic services that false-positive screen results would cause should not be overlooked. We, therefore, sought to derive a NGRS that could inform risk-stratified screening with greater efficiency than eligibility on age alone, using data routinely recorded in EHRs. Our analysis suggests that 5.6% more unnecessary screens would be undertaken to potentially detect the same proportion of cancers screened within the 50–74 population. Age at invitation for those with CRC ranged from 33–87-years.

Strengths and limitations. We acknowledge several limitations to the present study. First are the many biases associated with self-reported measures [58]. Biases due to social desirability, for example, were apparent for self-reported alcohol use and smoking habits, which were disproportionately missing relative to other items on the same questionnaire. Imputation of missing data also adds some degree of uncertainty to the model estimates.

Furthermore, self-reporting of past diagnoses may be subject to recall biases dependent on the nature of the morbidity [59] and, if systematic, could skew regression weights.

Second is the validity of our self-reported measures of lifestyle choices. Women were asked to approximate their weekly consumption of alcohol, which is highly likely to vary week-by-week and year-by-year. Furthermore, for ease of self-reporting, the unit asked for was the number of drinks per week, which owing to variances in the volume or proof is not a standard measure of the alcohol unit equivalent to 10 mL of pure alcohol. An alternative would be to encode alcohol use, as we did for Smoking, as ever or never drinkers—with the caveat that dose-dependent risk would not be factored into risk predictions [27].

Thirdly, we observed only 41% of the expected incidences, which suggests that the UKCTOCS cohort is healthier than the general population from which it was drawn from [60]. Bias toward selection of those with higher education and healthier proclivities could reduce the variation in covariate measures in disproportion to the national population, and could, therefore, reduce the power to detect small associations in undersampled covariates [61].

Additionally, the NGRS was derived from females of predominantly Caucasian descent, and therefore the reproducibility of our findings in other settings, including to the national population, is unknown. Nonetheless, women of average-risk tend to attain the risk of a 50-year man four years later at age 54 [62], therefore suggesting that risk prediction may benefit from gender-specific baseline hazards, as many current models have done (QCancer 10, for example [63]). Finally, the inclusion of pre-randomisation cancers diagnosed before the age of fifty requires us to rely on cancer registrations during a period when coverage was not as comprehensive as it is now [64] and also assumes that self-reported measures of Alcohol use and Smoking habit did not alter in light of their diagnoses.

Weaknesses notwithstanding, most CRC diagnoses were post-randomisation events informed by cancer & death registries in adjunct with secondary care records that in Chapter 3 we verified in comparison to clinical confirmation to be reliable for use in research. Additionally, predictive covariates were preselected on robust evidence reported in meta-analyses, which circumvented the need for contentious statistical selection methods. We also attempted to limit model overoptimism by classifying the risk of a bootstrap resample. Bootstrapping has greater efficiency over subsampling data into

training and validation sets and also has the advantage of reducing overoptimism in comparison to internal validation.

Findings in relation to existing literature. Our findings suggest that women in the top 10% of non-genetic risk have three-fold increased hazards to the geometric mean, and those in the top 1% have a 64-fold increased risk. Comparatively, when stratified with a polygenic risk score of 37 known SNP variants, men and women in the top percentile and decile of risk had a 2.9 and 1.8 increased risk over the population median [7]. The steeper risk gradient afforded by the NGRS may be reflective of the small yet additive effect SNPs have on risk (OR of most penetrant of 37 SNPs is 1.53) or on the higher risk component attributed to environmental factors. It is estimated, for example, that obesity and alcohol & tobacco use is attributed with almost as much of the risk component (24%) to that of heritable factors (35%) [10,65]. It may also be that the NGRS is poorly calibrated at the apex of risk, which can be evaluated by plotting the expected risk against the observed risk.

Covariates were preselected *a priori* on conclusive evidence of the risk modification as reported in meta-analyses. Ignoring for the time being the confidence intervals which were not incorporated into risk predictions, the regression weights for seven covariates (Arthritis, Benign, BMI, Diabetes, Gravidity, Height, Smoking) were concordant with meta-analyses, while ten (Alcohol, Cholesterol, Ethnicity (Chinese only), HRT, Hypertension, Hysterectomy, OCP, OCP years, Osteoporosis, Smoking) were discordant in the direction of risk modification.

Missingness was high for self-reported Alcohol use (30%) Cholesterol (25%), Hypertension (25%), Osteoporosis (25%), and Smoking habit (29%) and therefore discordant estimates may, at least in part, be attributed to the uncertainties of imputation. Alternatively, while high total blood cholesterol is associated with increased risk, this risk seems to be due to the effect of low-density (RR 1.04 (95% CI 0.60–1.81)) rather than high-density lipoproteins (RR 0.84 (95% 0.69–1.02), though is not definitive [43]. In clinical practice, the International Classification of Diseases (ICD) used by many EHRs have separate clinical codes for high high-density lipoprotein (E.78.0) and a deficiency in low-density lipoproteins (E78.6). Similarly, we also did not distinguish HRT involving oestrogen from oestrogen and progesterone, although both treatments are associated with a significant reduction in risk which we did not find. A potential explanation could be a skewed distribution of HRT duration since the risk reduction is stronger for users of ≥

10 years [66]. Self-reported OCP duration, although questionable with a mean of 4.1 years, could explain why we did not observe a risk modification from OCP use, with a meta-analysis showing a risk reduction after a minimum of five-years use [26]. On ethnicity, our model suggests that women of Chinese ancestry have an increased risk relative to Caucasians, while national data shows that rates of CRC incidence are 32% lower among Chinese living in the UK [23]. Given that the Ethnicity composition of the training set included few ethnic minorities, it may be prudent to restrict risk predictions to those from European descent or modify regression weights with robust estimates whereby necessary.

Implications for clinicians and policymakers. The study suggests how the provision of screening resources could be streamlined using data routinely recorded by EHRs. Nevertheless, screening eligibility informed by the current model would require more screens to be undertaken in order to potentially detect the same proportion of CRCs at present. Thus, the current evidence suggests that the NGRS would be no more efficient than current eligibility based on age alone. Alternatively, stakeholders may consider the use of the NGRS for identifying those at the apex of risk. A panel of ten SNPs, for example, is sufficient to identify the top 0.4% at highest risk whom should receive colonoscopy over gFOBT [6]. Pending further investigation, the NGRS could potentially triage admittance to the current high-risk surveillance programme, which already includes those with a previous benign polyp or IBD diagnosis [67].

Age at invitation for those triaged by the NGRS and who had CRC ranged from 33–87-years. The NGRS may thus have use in the detection of CRCs in young adults aged < 50-years who are currently ineligible for screening, which—at least in theory—are rapidly growing cancers for whom the patients stand to benefit from considerable gains in life years saved. Conversely, invitation of those older than 74 years is unlikely to prolong life or be cost-effective, and could paradoxically cause more harm owing to the side effects of treatments or to mental wellbeing. It may be thus sensible to limit the upper age of entry to 74 in alignment with the current BCSP.

In engineering covariates self-reported at both recruitment to the UKCTOCS and then by postal questionnaire at 3–5 (FUQ I) and/or 9–13-years (FUQ II) post randomisation—namely for Arthritis, Cholesterol, Diabetes, HRT, Hypertension, Hysterectomy, Osteoporosis—we took a simple and non-nuanced approach in which positivity was defined as self-reporting at any point of the study period. Doing so, however, overlooks the length of exposure to covariates, and assumes, for instance, the

same risk increase for a women affected by diabetes many years prior to recruitment aged 50 to another who was 74 at recruitment but whom did develop diabetes until much later. Doing so also fails to consider the number of subjects who successfully manage their lifestyle so that diabetes no longer posed a risk to developing CRC. The use of cross-sectional measurements (self-reported at the age of 50–74 years) that tend to increase with age—BMI and Waist circumference, for example—also presents similar challenges for risk-prediction. These are but few of the many nuances of time-varying covariates that has implications should a NGRS informed by EHRs be deployed for screening eligibility.

Unanswered questions and future research. While this study demonstrates in principle how EHRs could inform risk-stratified CRC screening, assertions cannot be made on how best a NGRS may be implemented in a screening programme, how it may inform the frequency or modality of screening, and how might it perform within the elderly for whom overdiagnosis is an issue. Future work should, first and foremost, discern the model's ability to consistently score screenees with potentially detectable CRC higher than those who may be exposed to unnecessary and potentially harmful screens, while also ensuring that the predicted risks align with the observed risks across the whole risk gradient [68]. Obtaining these metrics would allow for our model to be contextualised in relation to current risk-prediction models previously reviewed [69–71] and evaluated [72]. To aid in the interpretation of risk, it would be worthwhile transitioning to an absolute risk scale, which could be achieved with a parametric survival analysis able to calculate the baseline hazard function. Risk predictions would then also benefit from the inclusion of the time-related component—and thus Age—into the NGRS, which is currently based on only the covariate-related component (the linear predictor). Finally, our study does not suggest how risk-stratified screening will perform on the broader population. Ideally, risk stratification would be demarcated at population-defined threshold equivalent to the absolute risk a woman on first invitation to the screening, as has been done in studies on polygenic risk scores for breast, prostate [73], and colorectal cancers [7].

4.5 Conclusions

Risk-stratified screening based on a NGRS derived from covariates routinely recorded in EHRs is less efficient than eligibility based on age alone. Future work should refine and validate the NGRS in context of its performance relative to existing risk prediction models reviewed in the literature.

4.6 References

[1] Office for National Statistics, Cancer survival by stage at diagnosis for England (experimental statistics): adults diagnosed 2012, 2013 and 2014 and followed up to 2015, 2016.

[2] F.A. Macrae, K.G. Tan, C.B. Williams, Towards safer colonoscopy: a report on the complications of 5000 diagnostic or therapeutic colonoscopies, Gut. 24 (1983) 376–83.

[3] M.H. Robinson, J.D. Hardcastle, S.M. Moss, S.S. Amar, J.O. Chamberlain, N.C. Armitage, J.H. Scholefield, C.M. Mangham, The risks of screening: data from the Nottingham randomised controlled trial of faecal occult blood screening for colorectal cancer, Gut. 45 (1999) 588–92.

[4] D. Easton, N. Pashayan, P. Pharoah, Risk-stratified cancer screening, 2017.

[5] UK National Screening Committee, Note of the meeting held on the 29 June 2018 in Edinburgh, Scotland, 2018.

[6] M.G. Dunlop, A. Tenesa, S.M. Farrington, S. Ballereau, D.H. Brewster, T. Koessler, P. Pharoah, C. Schafmayer, J. Hampe, H. Völzke, J. Chang-Claude, M. Hoffmeister, H. Brenner, S. von Holst, S. Picelli, A. Lindblom, M.A. Jenkins, J.L. Hopper, G. Casey, D. Duggan, P.A. Newcomb, A. Abulí, X. Bessa, C. Ruiz-Ponte, S. Castellví-Bel, I. Niittymäki, S. Tuupanen, A. Karhu, L. Aaltonen, B. Zanke, T. Hudson, S. Gallinger, E. Barclay, L. Martin, M. Gorman, L. Carvajal-Carmona, A. Walther, D. Kerr, S. Lubbe, P. Broderick, I. Chandler, A. Pittman, S. Penegar, H. Campbell, I. Tomlinson, R.S. Houlston, Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42 103 individuals, Gut. 62 (2013) 871–881. doi:10.1136/gutjnl-2011-300537.

[7] M.J.E. Frampton, P. Law, K. Litchfield, E.J. Morris, D. Kerr, C. Turnbull, I.P. Tomlinson, R.S. Houlston, Implications of polygenic risk for personalised colorectal cancer screening, Ann. Oncol. 27 (2016) 429–434. doi:10.1093/annonc/mdv540.

[8] J.A. Cooper, N. Parsons, C. Stinton, C. Mathews, S. Smith, S.P. Halloran, S. Moss, S. Taylor-Phillips, Risk-adjusted colorectal cancer screening using the FIT and routine screening data: development of a risk prediction model, Br. J. Cancer. 118 (2018) 285–293. doi:10.1038/bjc.2017.375.

[9] A.E. Hall, S. Chowdhury, N. Hallowell, N. Pashayan, T. Dent, P. Pharoah, H. Burton, Implementing risk-stratified screening for common cancers: a review of potential ethical, legal and social issues, J. Public Health (Bangkok). 36 (2014) 285–291. doi:10.1093/pubmed/fdt078.

[10] P. Lichtenstein, N. V Holm, P.K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, K. Hemminki, Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland, N Engl J Med. 343 (2000) 78–85. doi:10.1056/NEJM200007133430201.

[11] K.G.M. Moons, D.G. Altman, J.B. Reitsma, J.P.A. Ioannidis, P. Macaskill, E.W. Steyerberg, A.J. Vickers, D.F. Ransohoff, G.S. Collins, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration, Ann. Intern. Med. 162 (2015) W1. doi:10.7326/M14-0698.

[12] I.J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J.K. Kalsi, N.N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, D.N. Crump, S.K. Davies, A. Dawnay, S. Dobbs, G. Fletcher, J. Ford, K. Godfrey, R. Gunu, M. Habib, R. Hallett, J.

Herod, H. Jenkins, C. Karpinskyj, S. Leeson, S.J. Lewis, W.R. Liston, A. Lopes, T. Mould, J. Murdoch, D. Oram, D.J. Rabideau, K. Reynolds, I. Scott, M.W. Seif, A. Sharma, N. Singh, J. Taylor, F. Warburton, M. Widschwendter, K. Williamson, R. Woolas, L. Fallowfield, A.J. McGuire, S. Campbell, M. Parmar, S.J. Skates, Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial, Lancet. 387 (2016) 945–956. doi:10.1016/S0140-6736(15)01224-6.

[13] M. Spaan, A.W. van den Belt-Dusebout, C.W. Burger, F.E. van Leeuwen, R. Schats, C.B. Lambalk, M. Kortman, J.S.E. Laven, C.A.M. Jansen, L.A.J. van der Westerlaken, B.J. Cohlen, D.D.M. Braat, J.M.J. Smeenk, J.A. Land, F. van der Veen, J.L.H. Evers, M.M.E. van Rumste, Risk of Colorectal Cancer After Ovarian Stimulation for In Vitro Fertilization, Clin. Gastroenterol. Hepatol. 14 (2016) 729–737.e5. doi:10.1016/j.cgh.2015.12.018.

[14] C.-Y. Li, B. Song, Y.-Y. Wang, H. Meng, S.-B. Guo, L.-N. Liu, H.-C. Lv, Q.-J. Wu, Age at Menarche and Risk of Colorectal Cancer: a Meta-Analysis, PLoS One. 8 (2013) e65645. doi:10.1371/journal.pone.0065645.

[15] H.-B. Guan, Q.-J. Wu, T.-T. Gong, B. Lin, Y.-L. Wang, C.-X. Liu, Parity and risk of colorectal cancer: a dose-response meta-analysis of prospective studies, PLoS One. 8 (2013) e75279. doi:10.1371/journal.pone.0075279.

[16] S. Chen, H. Liu, J. Li, G. Yang, Risk of Gastric and Colorectal Cancer After Tamoxifen Use for Breast Cancer, J. Clin. Gastroenterol. 49 (2015) 666–674. doi:10.1097/MCG.0000000000000262.

[17] K. Gaitskell, K. Coffey, J. Green, K. Pirie, G.K. Reeves, A.A. Ahmed, I. Barnes, V. Beral, Tubal ligation and incidence of 26 site-specific cancers in the Million Women Study, Br. J. Cancer. 114 (2016) 1033–1037. doi:10.1038/bjc.2016.80.

[18] B. Boursi, K. Haynes, R. Mamtani, Y.-X. Yang, Thyroid dysfunction, thyroid hormone replacement and colorectal cancer risk, J. Natl. Cancer Inst. 107 (2015) djv084. doi:10.1093/jnci/djv084.

[19] Cancer Research UK, Bowel cancer incidence statistics [UK, 2015, ICD-10 C18–C20], (2018). http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence (accessed February 28, 2018).

[20] E. Fuller, J. Mindell, G. Prior, Health Survey for England 2015, London, 2016.

[21] H. Jayasekara, J.C. Reece, D.D. Buchanan, D.J. Ahnen, S. Parry, M.A. Jenkins, A.K. Win, Risk factors for metachronous colorectal cancer or polyp: a systematic review and meta-analysis, J. Gastroenterol. Hepatol. 32 (2017) 301–326. doi:10.1111/jgh.13476.

[22] Y. Ma, Y. Yang, F. Wang, P. Zhang, C. Shi, Y. Zou, H. Qin, Obesity and risk of colorectal cancer: a systematic review of prospective studies, PLoS One. 8 (2013) e53916. doi:10.1371/journal.pone.0053916.

[23] National Cancer Intelligence Network (NCIN) and Cancer Research UK, Cancer Incidence and Survival by Major Ethnic Group, England, 2002-2006, 2009.

[24] K.J. Wernli, Y. Wang, Y. Zheng, J.D. Potter, P.A. Newcomb, The Relationship between Gravidity and Parity and Colorectal Cancer Risk, J. Women's Heal. 18 (2009) 995–1001. doi:10.1089/jwh.2008.1068.

[25] J. Green, B.J. Cairns, D. Casabonne, F.L. Wright, G. Reeves, V. Beral, Million Women Study collaborators, Height and cancer incidence in the Million Women Study:

prospective cohort, and meta-analysis of prospective studies of height and total cancer risk, Lancet Oncol. 12 (2011) 785–794. doi:10.1016/S1470-2045(11)70154-1.

[26] C. Bosetti, F. Bravi, E. Negri, C. La Vecchia, Oral contraceptives and colorectal cancer risk: a systematic review and meta-analysis, Hum. Reprod. Update. 15 (2009) 489–498. doi:10.1093/humupd/dmp017.

[27] V. Fedirko, I. Tramacere, V. Bagnardi, M. Rota, L. Scotti, F. Islami, E. Negri, K. Straif, I. Romieu, C. La Vecchia, P. Boffetta, M. Jenab, Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies, Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. 22 (2011) 1958–72. doi:10.1093/annonc/mdq653.

[28] K.K.F. Tsoi, C.Y.Y. Pau, W.K.K. Wu, F.K.L. Chan, S. Griffiths, J.J.Y. Sung, Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies, Clin. Gastroenterol. Hepatol. 7 (2009) 682-688.e1–5. doi:10.1016/j.cgh.2009.02.016.

[29] C. Schneider, S.S. Jick, U. Bothner, C.R. Meier, Cancer risk in patients with chronic obstructive pulmonary disease, Pragmatic Obs. Res. 1 (2010) 15–23. doi:10.2147/POR.S13176.

[30] E.M. Chun, S.W. Kim, S.Y. Lim, Prevalence of colorectal adenomatous polyps in patients with chronic obstructive pulmonary disease, Int. J. Chron. Obstruct. Pulmon. Dis. 10 (2015) 955. doi:10.2147/COPD.S83341.

[31] H. Freisling, M. Arnold, I. Soerjomataram, M.G. O'Doherty, J.M. Ordóñez-Mena, C. Bamia, E. Kampman, M. Leitzmann, I. Romieu, F. Kee, K. Tsilidis, A. Tjønneland, A. Trichopoulou, P. Boffetta, V. Benetou, H.B.A. Bueno-de-Mesquita, J.M. Huerta, H. Brenner, T. Wilsgaard, M. Jenab, Comparison of general obesity and measures of body fat distribution in older adults in relation to cancer risk: meta-analysis of individual participant data of seven prospective cohorts in Europe, Br. J. Cancer. 116 (2017) 1486–1497. doi:10.1038/bjc.2017.106.

[32] M.W.M.D. Lutgens, M.G.H. van Oijen, G.J.M.G. van der Heijden, F.P. Vleggaar, P.D. Siersema, B. Oldenburg, Declining Risk of Colorectal Cancer in Inflammatory Bowel Disease, Inflamm. Bowel Dis. 19 (2013) 789–799. doi:10.1097/MIB.0b013e31828029c0.

[33] Y. Komaki, F. Komaki, D. Micic, A. Ido, A. Sakuraba, Risk of Colorectal Cancer in Chronic Kidney Disease, J. Clin. Gastroenterol. (2017) 1. doi:10.1097/MCG.0000000000000880.

[34] Y. Komaki, F. Komaki, D. Micic, A. Ido, A. Sakuraba, Risk of colorectal cancer in chronic liver diseases: a systematic review and meta-analysis, Gastrointest. Endosc. 86 (2017) 93–104.e5. doi:10.1016/j.gie.2016.12.009.

[35] P.M. Rothwell, M. Wilson, C.-E. Elwin, B. Norrving, A. Algra, C.P. Warlow, T.W. Meade, Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials, Lancet. 376 (2010) 1741–1750. doi:10.1016/S0140-6736(10)61543-7.

[36] T. Lytras, G. Nikolopoulos, S. Bonovas, Statins and the risk of colorectal cancer: an updated systematic review and meta-analysis of 40 studies, World J. Gastroenterol. 20 (2014) 1858–70. doi:10.3748/wjg.v20.i7.1858.

[37] Y. Dong, J. Zhou, Y. Zhu, L. Luo, T. He, H. Hu, H. Liu, Y. Zhang, D. Luo, S. Xu, L. Xu, J. Liu, J. Zhang, Z. Teng, Abdominal obesity and colorectal cancer risk: systematic review and meta-analysis of prospective studies, Biosci. Rep. 37 (2017) BSR20170945. doi:10.1042/BSR20170945.

[38] K.J. Lin, W.Y. Cheung, J.Y.-C. Lai, E.L. Giovannucci, The effect of estrogen vs. combined estrogen-progestogen therapy on the risk of colorectal cancer, Int. J. Cancer. 130 (2012) 419–30. doi:10.1002/ijc.26026.

[39] G. Luo, Y. Zhang, L. Wang, Y. Huang, Q. Yu, P. Guo, K. Li, Risk of colorectal cancer with hysterectomy and oophorectomy: a systematic review and meta-analysis, Int. J. Surg. 34 (2016) 88–95. doi:10.1016/j.ijsu.2016.08.518.

[40] T.A. Simon, A. Thompson, K.K. Gandhi, M.C. Hochberg, S. Suissa, Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis, Arthritis Res. Ther. 17 (2015) 212. doi:10.1186/s13075-015-0728-9.

[41] L. Wu, C. Yu, H. Jiang, J. Tang, H. Huang, J. Gao, X. Zhang, Diabetes Mellitus and the Occurrence of Colorectal Cancer: an Updated Meta-Analysis of Cohort Studies, Diabetes Technol. Ther. 15 (2013) 419–427. doi:10.1089/dia.2012.0263.

[42] Y.J. Hee, C.S. Bang, G.H. Baik, I.S. Shin, K.T. Suk, T.Y. Park, D.J. Kim, Association between ischemic heart disease and colorectal neoplasm: a systematic review and meta-analysis, Springerplus. 5 (2016) 1664. doi:10.1186/s40064-016-3349-0.

[43] X. Yao, Z. Tian, Dyslipidemia and colorectal cancer risk: a meta-analysis of prospective studies, Cancer Causes Control. 26 (2015) 257–268. doi:10.1007/s10552-014-0507-y.

[44] K. Esposito, P. Chiodini, A. Capuano, G. Bellastella, M.I. Maiorino, C. Rafaniello, D.B. Panagiotakos, D. Giugliano, Colorectal cancer association with metabolic syndrome and its components: a systematic review with meta-analysis, Endocrine. 44 (2013) 634–647. doi:10.1007/s12020-013-9939-5.

[45] Y. Zhang, D.T. Felson, R.C. Ellison, B.E. Kreger, A. Schatzkin, J.F. Dorgan, L.A. Cupples, D. Levy, D.P. Kiel, Bone mass and the risk of colon cancer among postmenopausal women: the Framingham study, Am. J. Epidemiol. 153 (2001) 31–7.

[46] J.U. Lim, J.M. Cha, J. Il Lee, K.R. Joo, H.P. Shin, J.J. Park, J.W. Jeon, Osteoporosis Is Associated With the Risk of Colorectal Adenoma in Women, Dis. Colon Rectum. 56 (2013) 169–174. doi:10.1097/DCR.0b013e31826f8338.

[47] J. Barnard, D. Rubin, Small-sample degrees of freedom with multiple imputation, Biometrika. 86 (1999) 948–955. doi:10.1093/biomet/86.4.948.

[48] I. Hertz-Picciotto, B. Rockhill, Validity and efficiency of approximation methods for tied survival times in Cox regression, Biometrics. 53 (1997) 1151–6. http://www.ncbi.nlm.nih.gov/pubmed/9333345 (accessed April 27, 2019).

[49] R Core Team, R: a language and environment for statistical computing, (2016). https://www.r-project.org/.

[50] M. Stevenson, T. Nunes, C. Heuer, J. Marshall, J. Sanchez, R. Thornton, J. Reiczigel, J. Robison-Cox, P. Sebastiani, P. Solymos, K. Yoshida, G. Jones, S. Pirikahu, S. Firestone, R. Kyle, J. Popp, M. Jay, epiR: tools for the Analysis of Epidemiological Data, (2017). https://cran.r-project.org/package=epiR.

[51] H. Wickham, ggplot2: elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.

[52] S. van Buuren, K. Groothuis-Oudshoorn, mice: multivariate Imputation by Chained Equations in R, J. Stat. Softw. 45 (2011) 1–67.

[53] A. Robitzsch, S. Grund, T. Henke, miceadds: some additional multiple imputation functions, especially for mice, (2018). https://cran.r-project.org/package=miceadds.

[54] F.E. Harrell, rms: regression Modeling Strategies, (2018). https://cran.r-project.org/package=rms.

[55] T.M. Therneau, A Package for Survival Analysis in S, (2015).

[56] H. Heberle, G.V. Meirelles, F.R. da Silva, G.P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams, BMC Bioinformatics. 16 (2015) 169. doi:10.1186/s12859-015-0611-3.

[57] J.K. Lee, E.G. Liles, S. Bent, T.R. Levin, D.A. Corley, Accuracy of Fecal Immunochemical Tests for Colorectal Cancer: systematic review and meta-analysis, Ann. Intern. Med. 160 (2014) 171–181. doi:10.7326/M13-1484.

[58] A. Althubaiti, Information bias in health research: definition, pitfalls, and adjustment methods, J. Multidiscip. Healthc. 9 (2016) 211–7. doi:10.2147/JMDH.S104807.

[59] J. Barber, S. Muller, T. Whitehurst, E. Hay, Measuring morbidity: self-report or health care records?, Fam. Pract. 27 (2010) 25–30. doi:10.1093/fampra/cmp098.

[60] M. Burnell, A. Gentry-Maharaj, A. Ryan, S. Apostolidou, M. Habib, J. Kalsi, S. Skates, M. Parmar, M.W. Seif, N.N. Amso, K. Godfrey, D. Oram, J. Herod, K. Williamson, H. Jenkins, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, S. Leeson, D. Cruickshank, S. Campbell, L. Fallowfield, I. Jacobs, U. Menon, Impact on mortality and cancer incidence rates of using random invitation from population registers for recruitment to trials, Trials. 12 (2011) 61. doi:10.1186/1745-6215-12-61.

[61] E.A. Struijk, A.M. May, J.W.J. Beulens, C.H. van Gils, E.M. Monninkhof, Y.T. van der Schouw, M.W.M. Verschuren, J.M.A. Boer, B.H. Bueno-de-Mesquita, P.H.M. Peeters, Mortality and cancer incidence in the EPIC-NL cohort: impact of the healthy volunteer effect, Eur. J. Public Health. 25 (2015) 144–149. doi:10.1093/eurpub/cku045.

[62] H. Brenner, M. Hoffmeister, V. Arndt, U. Haug, Gender differences in colorectal cancer: implications for age at initiation of screening, Br. J. Cancer. 96 (2007) 828–831. doi:10.1038/sj.bjc.6603628.

[63] J. Hippisley-Cox, C. Coupland, Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study, BMJ Open. 5 (2015) e007825–e007825. doi:10.1136/bmjopen-2015-007825.

[64] H.O. Dickinson, J.A. Salotti, P.J. Birch, M.M. Reid, A. Malcolm, L. Parker, How complete and accurate are cancer registrations notified by the National Health Service Central Register for England and Wales?, J. Epidemiol. Community Health. 55 (2001) 414–22.

[65] K.F. Brown, H. Rumgay, C. Dunlop, M. Ryan, F. Quartly, A. Cox, A. Deas, L. Elliss-Brookes, A. Gavin, L. Hounsome, D. Huws, N. Ormiston-Smith, J. Shelton, C. White, D.M. Parkin, The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015, Br. J. Cancer. 118 (2018) 1130–1141. doi:10.1038/s41416-018-0029-6.

[66] L.S. Mørch, Ø. Lidegaard, N. Keiding, E. Løkkegaard, S.K. Kjær, The influence of hormone therapies on colon and rectal cancer, Eur. J. Epidemiol. 31 (2016) 481–489. doi:10.1007/s10654-016-0116-z.

[67] S.R. Cairns, J.H. Scholefield, R.J. Steele, M.G. Dunlop, H.J.W. Thomas, G.D. Evans, J.A. Eaden, M.D. Rutter, W.P. Atkin, B.P. Saunders, A. Lucassen, P. Jenkins, P.D. Fairclough, C.R.J. Woodhouse, Guidelines for colorectal cancer screening and

surveillance in moderate and high risk groups (update from 2002), Gut. 59 (2010) 666–89. doi:10.1136/gut.2009.179804.

[68] E.W. Steyerberg, Y. Vergouwe, Towards better clinical prediction models: seven steps for development and an ABCD for validation, Eur. Heart J. 35 (2014) 1925–31. doi:10.1093/eurheartj/ehu207.

[69] A.K. Win, R.J. Macinnis, J.L. Hopper, M.A. Jenkins, Risk prediction models for colorectal cancer: a review, Cancer Epidemiol. Biomarkers Prev. 21 (2012) 398–410. doi:10.1158/1055-9965.EPI-11-0771.

[70] G.K. Ma, U. Ladabaum, Personalizing Colorectal Cancer Screening: a Systematic Review of Models to Predict Risk of Colorectal Neoplasia, Clin. Gastroenterol. Hepatol. 12 (2014) 1624–1634.e1. doi:10.1016/j.cgh.2014.01.042.

[71] J.A. Usher-Smith, F.M. Walter, J.D. Emery, A.K. Win, S.J. Griffin, Risk Prediction Models for Colorectal Cancer: a Systematic Review, Cancer Prev. Res. 9 (2016) 13–26. doi:10.1158/1940-6207.CAPR-15-0274.

[72] J.A. Usher-Smith, A. Harshfield, C.L. Saunders, S.J. Sharp, J. Emery, F.M. Walter, K. Muir, S.J. Griffin, External validation of risk prediction models for incident colorectal cancer using UK Biobank, Br. J. Cancer. 118 (2018) 750–759. doi:10.1038/bjc.2017.463.

[73] N. Pashayan, S.W. Duffy, S. Chowdhury, T. Dent, H. Burton, D.E. Neal, D.F. Easton, R. Eeles, P. Pharoah, Polygenic susceptibility to prostate and breast cancer: implications for personalised screening, Br. J. Cancer. 104 (2011) 1656–1663. doi:10.1038/bjc.2011.118.

Chapter 5. Evaluating the Use of Existing Tumour Markers for the Early Detection of Colorectal Cancer Using Longitudinal Preclinical Sera

A Retrospective Case-Control Study Nested Within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

The current chapter is a slightly modified and rewritten analysis of the data published with the addition of data for FAP, TIMP-1, and VEGFA [1]. Another difference between is the cut-off threshold applied: here we defined the threshold calculated as that equal to a 5% false positive rate in our control samples, whereas previously we used the cut-off thresholds recommended in the manufacturer's protocol and commonly used throughout clinical biochemistry laboratories. The study design was conceived with help from Dr John Timms, Dr Aleksandra Gentry-Maharaj, and Dr Wendy Alderton. Mr Richard Gunu undertook all measurements for CA125, CEA, and CYFRA21-1. Dr Evangelia-Ourania Fourkala, Dr Sophia Apostolidou, Dr Andy Ryan, and Dr Aleksandra Gentry-Maharaj was involved in obtaining clinician's confirmation. The peer-review article was written with advice from Dr Sophia Apostolidou, Dr Andy Ryan, Professor Ian Jacobs, Professor Usha Menon, Dr Wendy Alderton, Dr Aleksandra Gentry-Maharaj, and Dr John Timms. I personally undertook all analyses presented here, including the creation of all Figures and Tables, and drafted all thesis versions with guidance from Professor Usha Menon and Dr Aleksandra Gentry-Maharaj.

5.1 Background

Detecting occult blood by the oxidation of guaiac (gFOBT) has been the cornerstone of many first-generation CRC screening programmes worldwide [2,3]. However, despite affording a 25% reduction in mortality to users [4], only 54% of first-time invitees in England partook during 2006–09 [5], and this has decreased year-on-year since 2010 [6]. Accordingly, there is scope to extend the success of the BCSP by merely improving uptake to screening.

Varied screening modalities may appeal to a range of preferences and thus widen coverage [7–9]. In particular, a blood test could address a number of perceived barriers to uptake [10,11] while also being familiar in routine clinical practice. Indeed, studies have reported that average-risk patients with a history of refusing stool-based screening prefer

to undergo a blood test [8,12]. Liquid biopsies are therefore an attractive method for CRC screening.

Proteins are the biological endpoints of genetic and epigenetic phenomena and may be secreted or leaked into circulation where they can be sampled as a liquid biopsy. No protein markers, however, have been recommended for screening purposes [13]—including CEA, the conventional marker for CRC which is approved for monitoring tumour burden and recurrence of clinical disease. The diagnostic and screening performance of CEA is presented in Table 5–1. Many of these performance estimates, however, are assessed using blood drawn at time diagnosis, and are thus arguably overestimated as the tumour grows exponentially, an inflammatory response is triggered, and symptoms become notable.

Table 5–1: Diagnostic and screening performance of CEA.

Marker	Specificity	Sensitivity
CEA [14]	1.00	0.42
CEA [15]	1.00	0.22
CEA [16]$_1$	0.99	0.12
CEA[17]	0.95	0.56
CEA [18]	0.95	0.45
CEA [19]	0.95	0.44
CEA [20]	0.95	0.43
CEA [21]	0.95	0.43
CEA [22]	0.95	0.39
CEA [23]	0.95	0.39
CEA [24]	0.95	0.37
CEA [25]	0.95	0.34
CEA [26]$_2$	0.95	0.19

Samples predating diagnosis by $_1$ 2.7 and $_2$ 0.3 years on average.

Preclinical samples drawn from asymptomatic patients in advance of diagnosis provide the best means of elucidating screening performance [27,28]. These are obtained only from large, mature biobanks, which have not been available until recently. Just two studies have accrued preclinical samples necessary to report the true screening performance of CEA. The first procured 120 cross-sectional sera drawn 0.13–10.3 years (median 2.7) before the eventual diagnosis of CRC (64 Dukes A–B, 52 Dukes C–D, 4 undeclared) [16]. They estimated that CEA, at 99% specificity, was positive (> 5.6 ng mL^{-1}) in 12% of these samples. This sensitivity increased to 30% (27, 43, 10, 43 for Dukes A–D, inclusively) when restricted to samples drawn 0.2–2 years before diagnosis. In another cross-sectional

study, CEA at 95% specificity (threshold not stated) conferred a lead time of up to 7 months in 19% of 32 (17 early stage/15 advanced stage) preclinical cases [26]. The current evidence, therefore, suggests that CEA is unsuitable for screening for CRC but may prove useful when combined with other markers that could complement its limitations. However, owing to the scarcity of preclinical samples, it is unknown how other markers identified using clinical samples fare for detecting CRC in asymptomatic patients.

Using longitudinal preclinical serum drawn from women recruited to the UKCTOCS who had a verified diagnosis of CRC and validated, commercial immunoassays, the present retrospective longitudinal case-control study aimed to evaluate the performance of existing markers described in the literature to detect asymptomatic CRC as a function of time before clinical diagnoses.

5.2 Methods

Study design. The present population-based, retrospective case-control study was nested within the multimodal arm of the UKCTOCS involving 50,640 women who donated blood annually. Subjects were recruited to the UKCTOCS from NHS patient registers to one of thirteen regional UKCTOCS centres in England, Northern Ireland, and Wales during 2001–05. Women with a previous bilateral oophorectomy or ovarian malignancy, heightened risk of familial ovarian cancer, or an active non-ovarian malignancy were excluded from recruitment to the UKCTOCS [29]. Only those who consented to their involvement in secondary studies were eligible for the study.

Subjects. One-hundred subjects were studied (40 CRCs, 20 benign adenomas, and 40 controls who were cancer-free). The Power to detect discrimination in marker measurements between 40 cases and 40 controls was calculated for AUCs ranging from 0.5–1.0 at an α level of 0.05 (Figure 5–1). The current sample size has a β level of 0.8 for detecting AUCs $\geq$ 0.675.

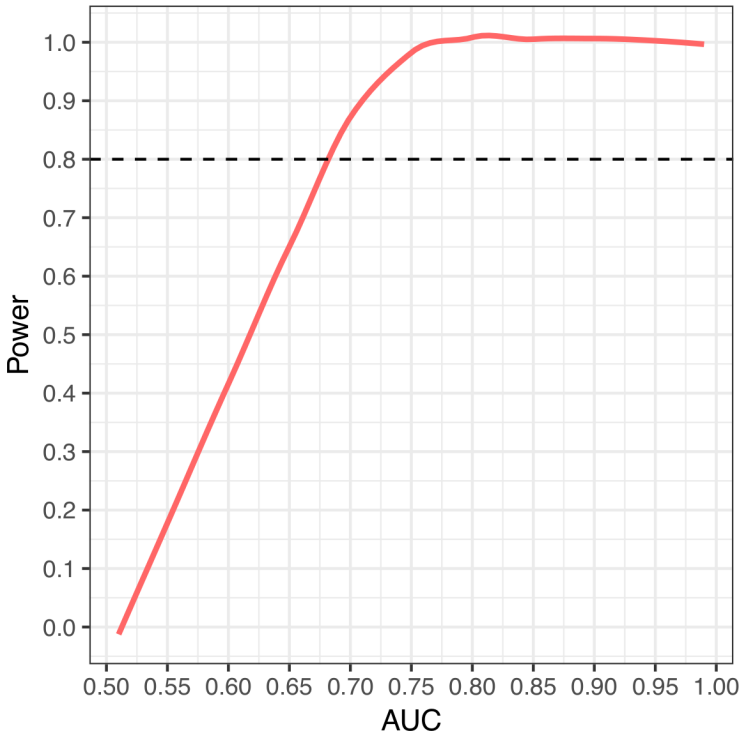


Figure 5–1: The loss of power for detecting increasingly indiscriminate markers. The likelihood of committing a type II error is ≥ 20% for markers with an AUC ≤ 0.675.

CRCs diagnosed post-randomisation to the UKCTOCS were informed via cancer & death registries, Hospital Episode Statistics (England only), self-reporting by postal questionnaire at 3–5 years after randomisation, or UKCTOCS trial resources and were verified by the patient's treating clinician in Chapter 3. Only women who gave ≥ three serial samples were eligible for selection (most were offered 7–11 screens). Women with neoplasms of the appendix (C18.1) or carcinoids and adenocarcinomas with signet-ring cell or mucinous morphology were excluded. Benign neoplasms of the colon and rectum (D12), excluding those of the appendix (D12.1), anus and anal canal (D12.9) were identified through HES (England only) for those without a cancer diagnosis of any origin according to CR & HES. Controls were defined as women with no diagnosis of benign neoplasms or carcinoma *in situs* (D010–12) according to CR & HES or having had no malignancy of any origin according to CR, DC, HES, and SR data. Definitions were applied on 1 September 2013. Subjects' baseline characteristics were obtained via a recruitment questionnaire (Supplementary Figure S1) completed by all subjects in-person at recruitment during 2001–05. These were BMI (derived from height & weight), Ethnicity, Gravidity (pregnancies lasting < 6 months), Parity (pregnancies lasting > 6 months), HRT use, OCP use, and previous hysterectomy. Missing data for HRT use was

supplemented with data self-reported via postal follow-up questionnaire I (Supplementary Figure S2) completed at 3–5-years post-randomisation.

Samples. The set consisted of 386 serum samples (154 controls, 78 benign, 154 cancers; Table 5–2) drawn from 100 subjects (40 controls, 20 beings, 40 cancers). Each woman had 3–4 samples eligible for the study (mean 3.9). Non-cancer controls were matched 1:1 to cancers by UKCTOCS trial centre, age at 0–1-year preclinical sample draw (± five years), and date of sample draw (same day for 0–1-year preclinical sample and ± four months for 3–4 years preclinical sample). Benigns were matched 1:1 with Dukes A–B cases by age at sample donation (± five years) and collection centre (excluding six cases in trial centres in Northern Ireland and Wales that were matched to the nearest trial centre in England). The lead time (years to diagnosis) was calculated as the years between the date of blood draw and verified date of diagnosis (controls took the date of diagnosis of their matched case). Time from draw to centrifugation was defined as the number of hours between blood draw and centrifugation as recorded in the UKCTOCS trial database.

Table 5–2: Sample counts according to lead time and clinical group.

Lead time (years)	Control	Benign	Dukes A–B	Dukes C–D
0–1	40	20	20	20
1–2	37	20	19	19
2–3	37	18	19	17
3–4	40	20	20	20
0–4	154	78	78	76

Marker selection. A literature review was undertaken to identify potential markers. Studies published until February 2013 were identified from Medline using the textword search "colorectal cancer AND detection OR diagnosis AND blood OR serum OR plasma AND marker OR biomarker" combined ("AND") with the Medical Subject Headings (MESH) search "Colorectal neoplasms". Results were restricted to full-text articles, including reviews, published in English. Studies were reviewed for their suitability, first by title and then by the abstract. Studies with a sample size < 20 were excluded as were markers with a specificity < 95%. Only protein markers were considered. Studies chanced upon were included until September 2013. The principle summary measures sought were sensitivity, specificity, and sample size. Assay type, cut-off thresholds, stage-specific sensitivities, and AUC were also extracted but were not necessary for inclusion. Eligible studies which 1) did not report the sensitivity for

detecting Dukes A–B cancers and 2) reported a Dukes A–B sensitivity < 35% were excluded. (This is not to suggest that 35% is a clinically relevant threshold—but is intended not to exclude all markers that may complement one another in a multi-marker panel). All markers satisfying the above criteria, and which had an available commercial assay kit, were selected. CEA was measured as the established marker for monitoring the burden or recurrence of CRC. CA125 was evaluated from trial measurements obtained for all UKCTOCS multimodal screens.

Immunoassays. Blood was drawn from UKCTOCS multimodal women by venepuncture and separated to serum according to a standardised protocol (Menon et al., 2009). From this, an aliquot was taken and assayed for CA125, with the remaining sera being dispensed in up to 10 straws of ~400 μL. Straws were heat-sealed and shipped frozen at the end of each day to a cryorepository for long-term storage at –196°C. For this study, samples (i.e. one straw per sample) were retrieved and shipped to the laboratory on dry ice, where they were pseudonymised, thawed at 4°C, randomly aliquoted into 96-well plates, and stored at –80°C until assayed. Assayed samples were identified by their pseudonymised ID and were relinked only after completion of all assays.

Serum CEA and CYFRA21-1 levels were determined on the Cobas e 411 platform (Roche Diagnostics) using the Elecsys CEA (cat no: 11820966122) and CYFRA21-1 (cat no: 11731629) assay kits. Elecsys CalSet and PreciControl tumour marker standards were used for quality control. CEA & CYFRA 21-1 determinations were carried out by a senior technician with extensive experience of the Cobas platform and who was unaware of the experimental design. CA125 measurements were those determined on the Cobas platform for the UKCTOCS multimodal arm. CA125, CEA, and CYFRA21-1 are presented as U mL-1, ng mL^{-1}, and ng mL^{-1}, respectively.

Commercial immunoassay kits were used to determine FAP (Thermo Scientific; cat no: EHFAP), TIMP-1 (R&D Systems; cat no: DTM100), and VEGFA (R&D Systems; cat no: DVE00) levels in serum. Assays were conducted according to the manufacturer's protocol. Standards ranged from 16.38–4,000 ng mL^{-1} (1:2.5 serial dilution) for FAP, 0.156–10 ng mL^{-1} (1:2 serial dilution) for TIMP-1, and 31.3–2,000 pg mL^{-1} (1:2 serial dilution for VEGFA. Samples were diluted 1:100 for FAP, 1:100 for TIMP-1, and were undiluted for VEGFA. Samples were measured once and standards twice. Absorbance measurements were normalised to that of a background reaction containing only buffer. Determinations were made by interpolation from a four-parametric logistic curve fitted

to standards measured on the same plate and corrected for dilution factor. FAP and TIMP-1 are presented as ng mL^{-1}. VEGFA is presented as pg mL^{-1}.

Data analyses

Subjects. Ethnicity was stratified as White, Black (Black-African/Black-Caribbean/Black-Other) or Other (Indian/Pakistani/Chinese/Bangladeshi/Other). BMI was calculated from self-reported height and weight. Differences in the baseline characteristics between clinical groups were assessed using multivariate analysis after listwise deletion of missing data (X^2 for categorical; Kruskal-Wallis & post hoc pairwise Mann-Whitney with Bonferroni correction).

Samples. Differences in time to centrifugation and years between menopause and blood draw between CRCs, benign adenomas, and controls were determined by ANOVA/Kruskal-Wallis with Tukey's post hoc test.

Immunoassay performance. Recovery (%) was calculated for each by comparing the determined and actual concentration of 6–8 standards. Recovery was not calculated for TIMP-1 due to insufficient assay wells remaining. The Cobas CA125, CEA, and CYFRA21-1 assays are validated for clinical use and their quality was assessed using internal quality controls (unreported).

Differentiation. Differences in marker levels between clinical groups were assessed using multivariable analysis (Kruskal-Wallis & post hoc pairwise Mann-Whitney with Bonferroni correction). Boxplots denoting the median, IQR, range, and outliers were plotted.

Classification. Cut-off thresholds were calculated at 95% specificity (i.e. a 5% false-positive rate) assuming both up and downregulation during cancer. The sensitivity (TP/(TP+FN)) of each marker were computed with the threshold yielding the highest AUC.

Longitudinal profiles. Average marker serial profiles were modelled with Loess regression. Loess (Locally Estimated Scatterplot Smoothing) regression is a non-linear extension of least-squares regression wherein a curve is modelled on many localised subsets of data. Lead time was estimated by interpolation of Loess curves at the threshold

equivalent to 95% specificity (CA125 < 7.27 U mL^{-1}, CEA > 5.13 ng mL^{-1}, $CYFRA21^{-1}$ > 2.87 ng mL^{-1}, FAP < 18.48 ng mL^{-1}, TIMP-1 < 28.55 ng mL^{-1}, VEGFA > 844.52 pg mL^{-1})**.** Within-person variation in markers levels across serial samples was assessed using the coefficient of variation (CV). Individual CVs were pooled into Dukes A–B & C–D CRC, benign adenomas, and controls for which an aggregate median and range were computed. Inter-group differences in individual CVs were assessed using multivariable analysis (Kruskal-Wallis & post hoc pairwise Wilcox rank sum test with Bonferroni adjustment).

Statistical analyses. All analyses were made with R version 3.3.2 [30] running drc [31], epiR [32], ggplot2 [33], Hmisc [34], pROC [35], OptimalCutpoints [36], and tidyr [37] packages. A p value < 0.05 was considered significant.

5.3 Results

Marker selection. The specified search terms retrieved 155 studies. After removing five duplicated studies, 113/150 studies were excluded after a title review, 30/37 after an abstract review, and 4/7 after a full-text review. Three studies [16,20,38] passed all selection criteria. The hit rate of the search strategy was 2% (3/150). Eligible studies were pooled with 13 studies chanced upon [14,15,18,19,22–24,39–44] for a total of 16 studies. Overall, 43 proteins (34 unique) were evaluated in the 16 studies. The majority of these assessed the performance of CEA in parallel. Twenty-four proteins (22 unique) were excluded for not reporting stage-specific sensitivities (Table 5–3).

Table 5–3: Diagnostic performance of literature markers excluded for not reporting stage-specific sensitivity.

Marker	Specificity	Sensitivity
AFP [19]	0.95 (0.95)	0.7 (0.44)
ASC [19]	0.95 (0.95)	0.18 (0.44)
CA125 [19]	0.95 (0.95)	0.18 (0.44)
CA15-3 [19]	0.95 (0.95)	0.4 (0.44)
CA19-9 [38]	0.95 (0.95)	0.18 (0.43)
CA19-9 [19]	0.95 (0.95)	0.23 (0.44)
CA72-4 [19]	0.95 (0.95)	0.21 (0.44)
$CA242_1$ [16]	0.92	0.10
CCSA2 [43]	1.00	0.15
DR70 [18]	0.95	0.7 (0.45)
HCGβ [19]	0.95 (0.95)	0.4 (0.44)

HGF [19]	0.95 (0.95)	0.30 (0.44)
IL6 [19]	0.95 (0.95)	0.31 (0.44)
IL8 [19]	0.95 (0.95)	0.28 (40.4)
MMP9 [18]	0.95	0.13 (0.45)
NNMT [19]	0.95 (0.95)	0.31 (0.44)
NNMT [22]	0.95 (0.95)	0.51 (0.39)
NSE [19]	0.95 (0.95)	0.14 (0.44)
PSME3 [19]	0.95 (0.95)	0.20 (0.44)
S100A8 [15]	0.95 (1.00)	0.41 (0.22)
S100A9 [15]	0.95 (1.00)	0.44 (0.22)
S100A12 [19]	0.95 (0.95)	0.20 (0.44)
sCD26 [18]	0.95 (0.95)	0.26 (0.45)
Serpin B5 [19]	0.95 (0.95)	0.24 (0.44)

[1] preclinical samples predating diagnosis by 2.7 years on average. Metrices within parenthesis () indicates the performance of CEA using the same samples.

Shortlisted proteins were ranked by Dukes A–B sensitivity (Table 5–4) and 14 proteins (8 unique) with a Dukes A–B sensitivity < 35% were excluded. The remaining proteins (CYFRA21-1, FAP, TIMP-1, VEGFA) had commercial assays available and were evaluated alongside CEA and CA125.

Table 5–4: Diagnostic performance of shortlisted proteins.

Marker	Specificity (%)	Sensitivity (%)				
		Dukes A–B	Dukes A	Dukes B	Dukes C	Dukes D
VEGFA [41]	95	54	37	61	63	86
TIMP-1 [23]	95 (95)	48	48		—	—
TIMP-1 [20]	95 (95)	42[1]	42[1] (30[1])		—	—
FAP [19]	95 (95)	44	36 (13)	49 (37)	51 (34)	32 (88)
CYFRA21-1 [24]	95 (95)	37	18 (0)	40 (33)	49 (46)	86 (71)
MMP2 [20]	95 (95)	33	33[1] (30[1])		—	—
CAM43 [14]	97 (100)	31	0 (10)	41 (34)	22 (18)	77 (77)
PSA (Free) [42]	97	28	28		9	
CA50 [44]	97	26	26		47	
Ferritin [19]	95 (95)	23	13 (13)	31 (37)	34 (34)	13 (88)
Anti-p53 [39]	100	22	20	22	22	25
CYFRA21-1 [19]	95 (95)	21	13 (13)	27 (37)	24 (34)	81 (88)
OPN [19]	95 (95)	18	13 (13)	22 (37)	28 (34)	54 (88)
Anti-p53 [19]	95 (95)	17	16 (13)	18 (37)	16 (34)	24 (88)
CA19-9 [44]	100	16	16		44	
TIMP-1 [40]	98	14	5	25	14	
CA19-9 [24]	95 (95)	12	9 (0)	13 (33)	41 (47)	43 (71)
CRP [40]	98	11	10	13	31	
MMP9 [20]	95 (95)	0	0[1] (30[1])		—	—

[1] Including carcinoma *in situs*. Values in parenthesis () indicate the performance of CEA using the same sample set.

Subjects. The baseline characteristics of subjects at recruitment to the UKCTOCS are presented in Table 5–5. There was no evidence of inter-group differences in the baseline characteristics obtained.

Table 5–5: Baseline characteristics.

Characteristic	Control (n 40)	Benign (n 20)	Cancer (n 40)	*p* value
BMI (Kg/m²)				0.387
Median	25.2	25.8	24.2	
IQR	5.4	6.2	4.2	
Missing (n)	1	0	0	
Pregnancies < 6 months (n)				0.557
Median	0	0	0	
IQR	0	1	1	
Missing (n)	0	0	0	
Pregnancies > 6 months (n)				0.637
Median	2	2	2	
IQR	2	1	1	
Missing (n)	0	0	0	
Ethnicity				0.305
White	39 (97.5%)	18 (90.0%)	38 (95.0%)	
Black	0 (0.0%)	2 (10.0%)	1 (2.5%)	
Other	1 (2.5%)	0 (0.0%)	1 (2.5%)	
Missing	0 (0.0%)	0 (0.0%)	0 (0.0%)	
HRT (n)				
Never	33 (82.5%)	15 (75.0%)	32 (80.0%)	0.791
Current	7 (17.5%)	5 (25.0%)	8 (20.0%)	
Missing	0 (0.0%)	0 (0.0%)	0 (0.0%)	
OCP (n)				0.313
Never	23 (57.5%)	8 (40.0%	24 (60.0%)	
Ever	17 (42.5%)	12 (60.0%)	16 (40.0%)	
Missing	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Hysterectomy (n)				
No	32 (80.0%)	14 (70.0%)	32 (80.0%)	0.627
Yes	8 (20.0%)	6 (30.0%)	8 (20.0%)	
Missing	0 (0.0%)	0 (0.0%)	0 (0.0%)	

p values calculated after listwise deletion of missing data. Abbreviations: IQR, interquartile range; BMI, Body Mass Index; HRT, Hormone Replacement Therapy; OCP, Oral Contraceptive Pill.

The clinical & histological traits of benign adenomas and CRCs are summarised in Table 5–6. There was no evidence of inter-group differences in the affected primary site (*p* 0.155) or histological grade of lesions (*p* 0.999).

Table 5–6: Clinical & histological characteristics of benign adenomas and CRCs.

	Benign (n 20)	Dukes A–B (n 20)	Dukes C–D (n 20)
Primary site			
Caecum	1 (5%)	3 (15%)	4 (20%)
Ascending colon	0 (0%)	6 (30%)	2 (10%)
Transverse colon	3 (15%)	0 (0%)	3 (15%)
Splenic flexure	0 (0%)	1 (5%)	0 (0%)
Descending colon	2 (10%)	1(5%)	0 (0%)
Sigmoid colon	8 (40%)	5 (25%)	5 (25%)
Rectosigmoid jun.	1 (5%)	0 (0%)	0 (0%)
Rectum	2 (10%)	3 (15%)	6 (30%)
Undetermined	3 (15%)	1 (5%)	0 (0%)
Dukes stage			
A	—	7 (35%)	0 (0%)
B	—	13 (65%)	0 (0%)
C	—	0 (0%)	14 (70%)
D	—	0 (0%)	6
Histological grade			
I	NA	1 (5%)	2 (10%)
II	NA	12 (60%)	12 (60%)
III	NA	3 (15%)	2 (10%)
Unknown	NA	4 (20%)	4 (20%)
Morphology			
Adenocarcinoma	0 (0%)	20 (100%)	20 (100%)

Samples. There was no evidence of inter-group differences in the time between sample draw and centrifugation (p 0.195) or years from menopause to blood draw (p 0.433).

Immunoassays. Measurements were missing for CEA (n 5), CYFRA21-1 (n 5), CA125 (n 5), and VEGFA (n 6) owing to serum unavailability. Determinations were made against four-parametric logistic curves ran over five 96-well plates (Figure 5–2). Five TIMP-1 determinations were extrapolated from absorbencies marginally greater than the highest standard (> 10 ng mL^{-1}). These were determined as (after correction for 1:10 dilution) 1011.28 (SE 164.60), 1092.18 (181.76), 1147.06 (193.57), 1247.01 (215.40), and 1442.36 (259.17) ng TIMP-1 mL^{-1}. Ten determinations that had absorbance values lower than the lowest standard (< 15.6 ng mL^{-1} after correction for dilution) were taken at face value.

The performance of assay kits was assessed for those with surplus wells available. Recoveries for 4,000, 1,600, 256, 40.96, and 16.38 pg FAP mL^{-1} were 87% (3479/4000), 106% (1693/1600), 100% (256/256), 123% (50/40.96), and 47% (8/16.38), respectively.

Unfortified buffer (0 pg FAP mL^{-1}) was correctly determined as 0 (or < limit of quantitation). Recoveries for 2,000, 1,000, 500 250 125, 62.5, and 31.3 pg VEGFA mL^{-1} were 100% (1994/2000), 105% (1050/1000), 101% (505/500), 114% (143/125), 113% (71/62.5), and 98% (31/31.3), respectively. Unfortified buffer (0 pg VEGFA mL^{-1}) was correctly determined as 0 (or < limit of quantitation).

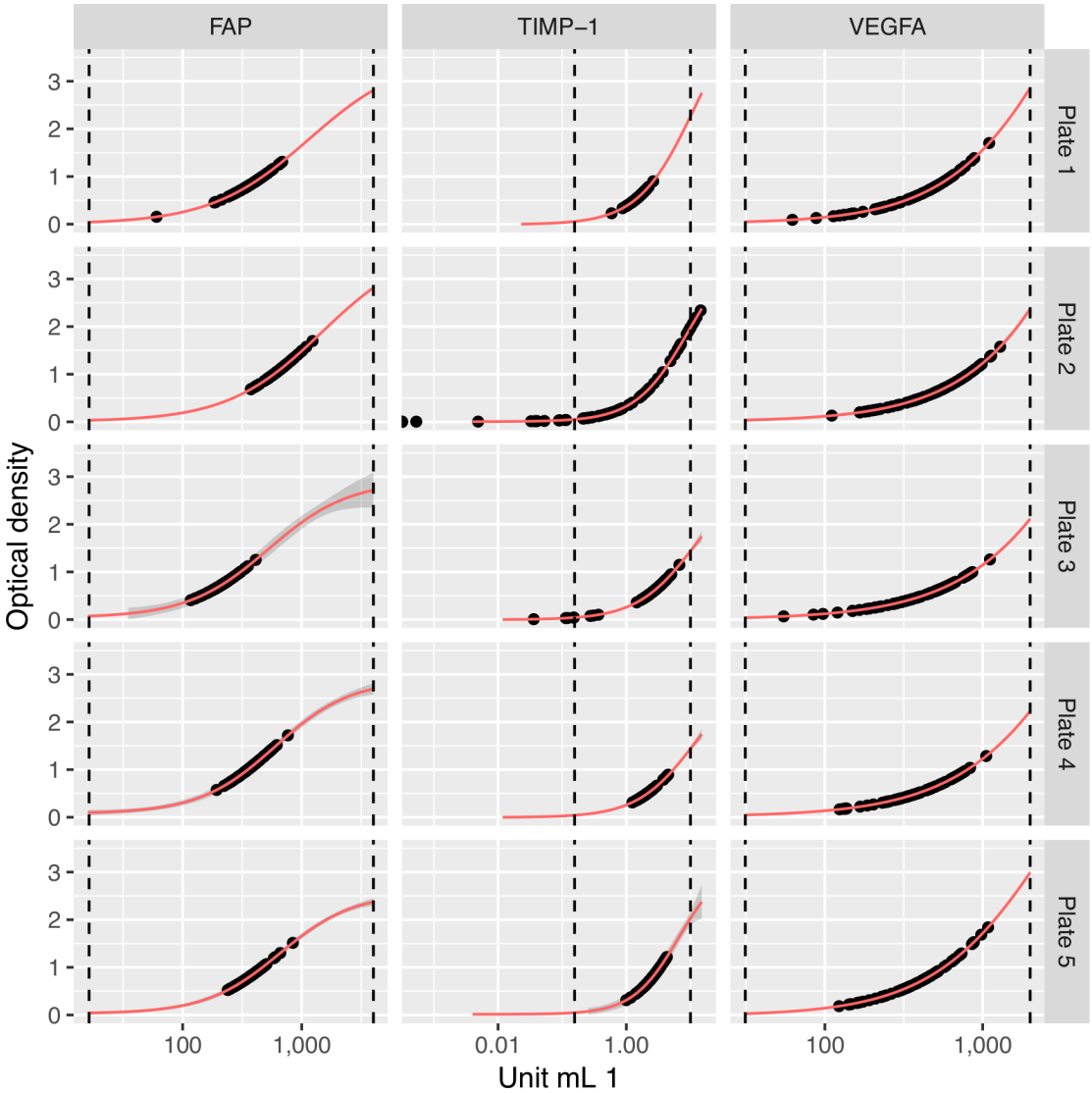


Figure 5–2: Four-parametric logistic curves (red), interpolated samples (black), and 95% confidence intervals (grey). Dashed lines demarcate upper and lower calibration standards. FAP and TIMP-1 are expressed as ng mL-1 and VEGFA as pg mL-1.

Differentiation. An overview of levels of each marker as a function of time before the diagnoses of benign adenoma or CRC and corresponding levels in women who remained cancer-free in is presented in Figure 5–3. Notably, CEA was upregulated in preclinical cancer samples relative to controls (2.05 vs 1.53 ng mL^{-1}; *p* 3.7e-05) and benign samples (2.05 vs 1.51 ng mL^{-1}; *p* 2.2e-04). Relative to controls, CA125 was significantly

downregulated in benign (p 0.011) but not cancer samples. No other comparisons were significant at $p < 0.05$.

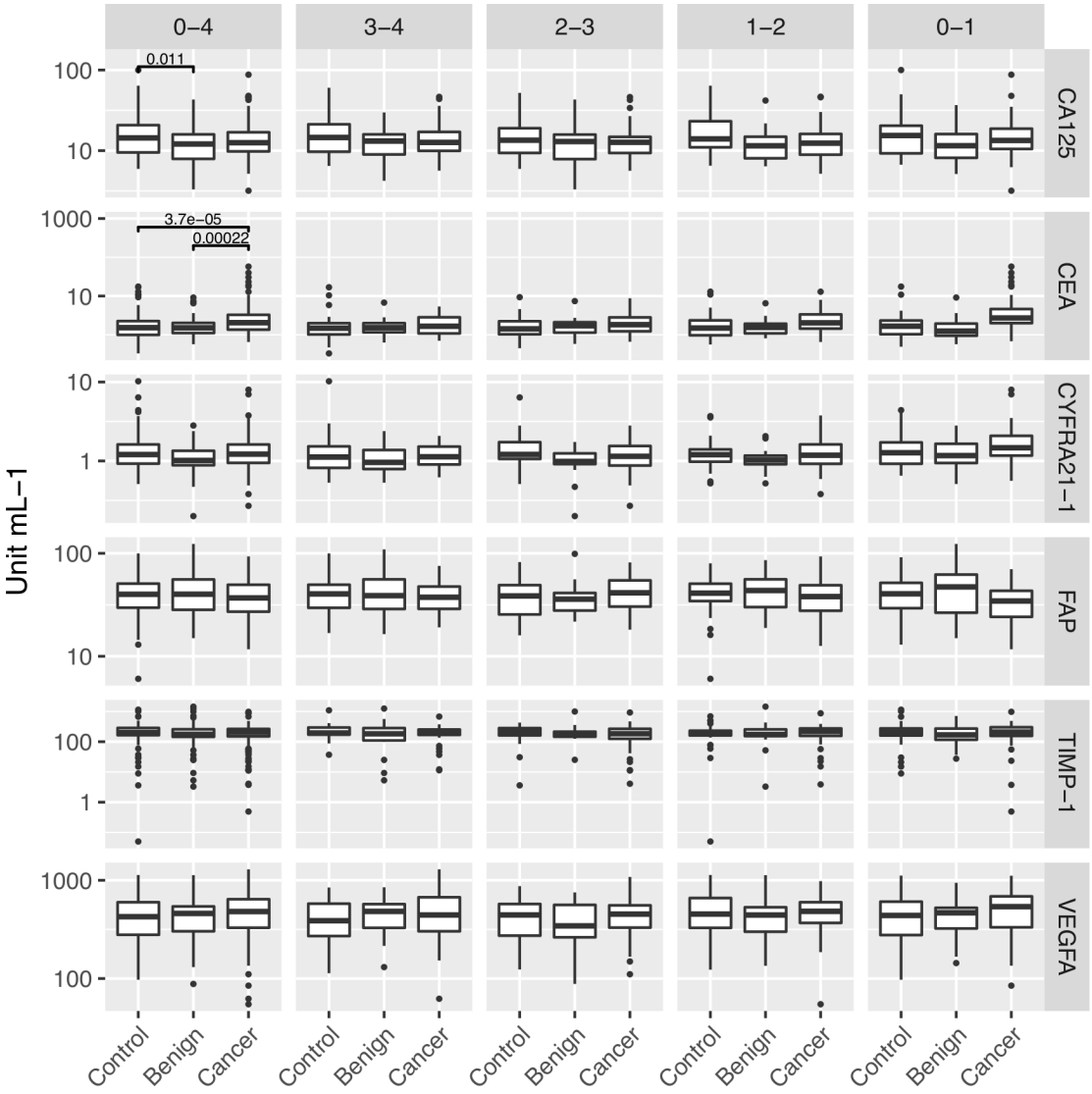


Figure 5–3: CEA is significantly elevated 0–4 years before cancer diagnosis. Boxplots denote the 25/50/75th percentiles, range, and outliers. Time to diagnosis for controls corresponds with that of their matched case. CA125 expressed as Units mL^{-1}, CEA, CYFRA21-1, FAP, and TIMP-1 as ng mL^{-1}, and VEGFA as pg mL^{-1}.

It was then sought if CEA was raised uniformly across all clinical stages at diagnosis. Median serum CEA for the 0–1, 1–2, 2–3, and 3–4 years preclinical groups were 2.61, 2.51, 2.40, and 2.27 ng mL^{-1} for Dukes A–B and 2.81, 1.65, 1.26, and 1.17 ng mL^{-1} for Dukes C–D CRCs, respectively (Table 5–7). CEA levels were significantly higher in Dukes A–B cancers at 2–3 and 3–4 years before diagnosis, but not at 0–1 and 1–2 years.

Table 5–7: CEA (ng mL^{-1}) levels in sera drawn in the years before diagnosis of Dukes A–B and C–D cancers.			
Years lead time	**Dukes A–B**	**Dukes C–D**	***p* value**
0–1	2.61	2.81	0.829
1–2	2.51	1.65	0.133
2–3	2.40	1.26	0.007*
3–4	2.27	1.17	0.002*
0–4	2.49	1.64	0.001*

Classification. AUCs were calculated assuming both down- and upregulation for all marker determinations at a threshold equivalent to 95% specificity. Performances were thereafter determined using the direction of dysregulation giving the highest AUC. Accordingly, CA125, FAP, and TIMP-1 were downregulated in preclinical cancer, while CEA, CYFRA21-1, and VEGFA were upregulated. The thresholds applied are outlined in Table 5–8.

Table 5–8: Cut-off thresholds equivalent to 95% specificity.		
Marker	**Dysregulation**	**Threshold**
CA125	<	7.27
CEA	>	5.13
CYFRA21-1	>	2.87
FAP	<	18.48
TIMP-1	<	28.63
VEGFA	>	844.52
CA125 presented as Units mL^{-1}, CEA, CYFRA21-1, TIMP-1, and FAP as ng mL^{-1} and VEGFA as pg mL^{-1}.		

The sensitivity of each marker for detecting preclinical benign adenoma and CRC as a function of time before diagnosis is presented in Figure 5–4. The sensitivity for detecting cancer predating diagnoses by 0–4 years ranked in order of CEA (12%), TIMP-1 (10%), VEGFA (9%), CA125 (7%), FAP (5%), and CYFRA21-1 (4%). The sensitivities for benign cases ranked in order of CA125 (14%), TIMP-1 (10%), CEA (5%), FAP (4%), VEGFA (4%), and CYFRA21-1 (0%). CEA had the greatest sensitivity for detecting cancers up to one year before diagnosis (23%), followed jointly by CYFRA21-1, TIMP-1, and VEGFA (12%), FAP (7%), and CA125 (5%). CA125 had a greater performance for detecting benign adenomas over cancers. Overall CEA positivity (5%) for benign adenomas was equal to that of the controls (5%). AUCs ranked in order of CEA (0.645), CA125 (0.563), VEGFA (0.554), FAP (0.545), CYFRA (0.512), and TIMP-1 (0.509).

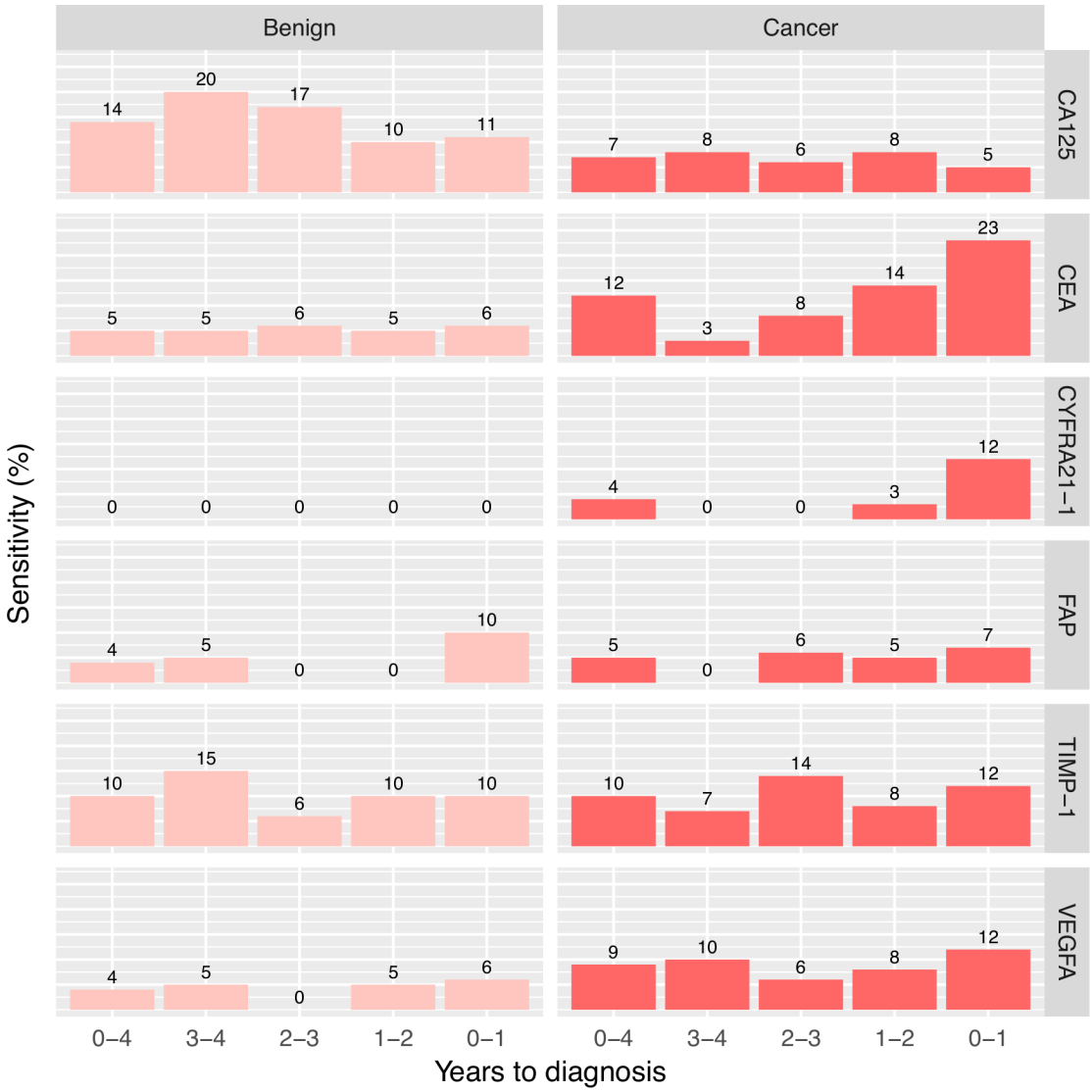


Figure 5–4: Sensitivity of markers for detecting preclinical benign adenoma and colorectal cancer as a function of time before clinical diagnosis at a threshold equivalent to 95% specificity.

Longitudinal profiles. The complexity of individual serial marker profiles is best illustrated in Figure 5–5, which denotes both individual and average profiles for each marker. Average CEA levels exceeded the threshold equivalent to 95% specificity at a lead time of 0.8 years. All other loess curves did not breach predetermined cut-off thresholds, though CYFRA21-1 is seemingly upregulated toward diagnosis.

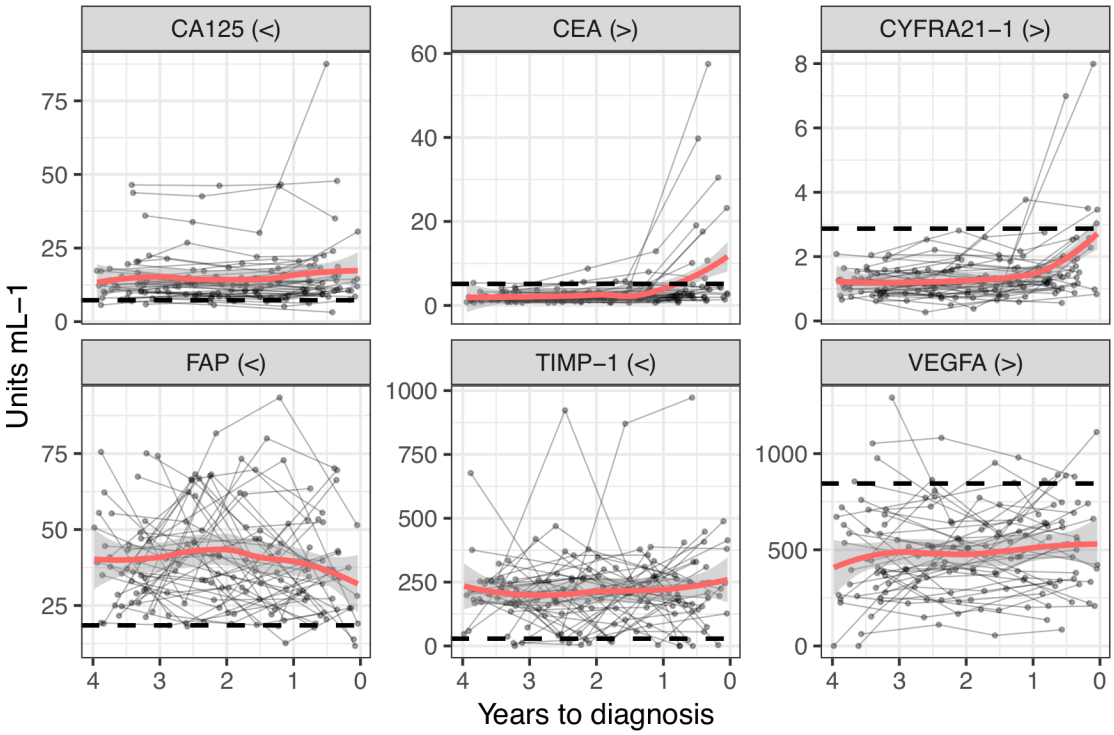


Figure 5–5: Marker levels in the lead-up to diagnosis of Dukes A–D cancers and trend profiles based on loess regression (red). Direction of dysregulation (< or >) noted in facet titles. Cut-off thresholds equal to 95% specificity demarcated with the dashed line.

It was then sought if each had their temporal baseline profile and whether deviation from this is indicative of preclinical cancer (Table 5–9). Individual CVs in CEA levels for future Dukes C–D patients (38% CV) significantly differed from Dukes A–B (11% CV; p 0.048), benign (13% CV; p 0.028), and controls (15% CV; p 0.003). All other comparisons for all markers were insignificant.

Table 5–9: Median and range in the variation of individual marker longitudinal profiles.

Marker	Control (n 40)	Benign (n 20)	Dukes A–B (n 20)	Dukes C-D (n 20)
CA125	10 (1–31)	11 (5–46)	12 (4–37)	12 (2–58)
CEA	15 (5–37)	13 (6–58)	11 (3–126)	38 (3–171)*
CYFRA21-1	19 (3–86)	24 (3–57)	20 (9–89)	29 (7–123)
FAP	34 (5–78)	35 (17–53)	35 (9–58)	32 (7–70)
TIMP-1	40 (9–116)	68 (16–106)	57 (12–118)	63 (10–111)
VEGFA	19 (6–64)	21 (5–43)	24 (6–75)	15 (7–39)

Variation described as CVs (%).

5.4 Discussion

Principle findings. Patients who prefer blood over stool-based screening commonly cite reasons of ease and convenience for their choice [8,12]. The present study, therefore, evaluated the screening performance of promising serum markers. Neither CA125, CEA, CYFRA21-1, FAP, TIMP-1 nor VEGFA is suitable for detecting preclinical CRC in asymptomatic patients and were all inferior to the gFOBT. Inadequacies notwithstanding, CEA was elevated in a subset of women up to two years before diagnosis.

Strengths and limitations. The findings of the present study depend on the robustness of marker determinations. Accordingly, CA125, CEA, and CYFRA21-1 were measured with assays marketed for *in vitro* diagnostic use. Indeed, the clinical performance of the CA125 and CEA assays have been verified [45]. The performance of assays selected for FAP, TIMP-1, and VEGFA, while intended for research only, were tested by the manufacturers and, where surplus wells were available, recovery of quality controls in the magnitude of serum concentrations was accurate. Another key strength is that sera were drawn and processed under standard operating procedures and were re-requested on failure of quality control. Further still, subjects were matched according to collection centre and date of blood draw in an attempt to minimise confounders in sample handling. Finally, the present study fulfils the criteria for the Prospective-specimen-collection, retrospective-blinded-evaluation (PRoBE) required to evaluate a marker's utility [28].

One should also be mindful of the inherent weaknesses, however. A significant limitation is that only sera drawn from postmenopausal women were studied—and may, therefore, lack application to the general population, including men who are more likely to be afflicted with CRC [46]. Secondly, high-risk benign adenomas of ≥ 1 cm or tubulovillous or villous histology most likely to progress to malignancy [47] were not distinguished from those unlikely to cause harm. Next, the literature search strategy, while defined, was not systematic and retrieved few studies reporting screening performance using preclinical samples. A systematic review could review the literature stored in Embase, Web of Science, and bioRxiv. Moreover, the addition of the terms 'screening', 'asymptomatic', and 'preclinical' to the textword search may have yielded more precise results. Finally, the present study is relatively small. If one wanted to evaluate promising markers on a larger yet different sample set, the UKCTOCS biobank holds many cross-sectional samples drawn at recruitment from the ~152,000 women randomised to the

ultrasound and control arms; some of whom have since been verified to have been diagnosed with CRC.

Findings in relation to existing literature. In two previous studies that studied cross-sectional PRoBE-compliant samples, CEA was 12% sensitive at leads times up to 10.3 years [16] and 19% sensitive up to 7 months before diagnosis [26]. Herein, CEA was 23%, 14%, 8%, and 3% sensitive at lead times ranging from 0–1, 1–2, 2–3, and 3–4 years, respectively. Given the inverse relationship between the sensitivity and lead time afforded, it is likely then that CEA levels correspond with tumour burden, and that only in a small subset of preclinical CRCs is it expressed at levels exceeding the clinical threshold. Variance in sensitivity estimates reported here and in the literature is thus likely due to differences in the lead-time composition of samples. Another notable difference between studies is the false-positive rate used to determine cut-off thresholds; ours and that of Ladd evaluated performance at 95% specificity, whereas Palmqvist opted for 99% specificity. Only through a hypothetical scenario that factors for disease prevalence in an average-risk population can an acceptable false-positive rate be determined.

CYFRA21-1, FAP, TIMP-1, and VEGFA were identified from the literature as promising diagnostic markers, and in many instances outperformed CEA with sensitivities for Dukes A–B cancers of ≥ 37%. None of these, however, performed as well as CEA and, expectantly, were all inferior for detecting preclinical disease. CYFRA21-1, TIMP-1, and VEGFA, for example, had sensitivities of 12% at a lead time ≤ one year, while FAP had a sensitivity of 7%. Perhaps unsurprisingly, CA125 had a sensitivity of 7%. These studies, like ours, used commercial immunoassay kits for blood measurements with exception for FAP [19] and TIMP-1 [23], which were measured using assays developed and validated in-house. Unlike ours, however, the diagnostic performance of TIMP-1 [20,23] and VEGFA [41] was determined via plasma measurements, which owing to their release from platelets during serum coagulation may prevent direct comparisons [48,49].

It may reassuring that many of the markers are implicated in carcinogenesis: CEA in altering cell adhesion [50]; FAP during the remodelling of the extracellular matrix [51]; TIMP-1 in creating an auspicious microenvironment [52]; and VEGFA in developing the vasculature needed to nourish a developing tumour [53]. Indirectly, CYFRA21-1—a fragment of cytokeratin 19—is a marker of cellular proliferation and apoptosis [54]. An exhaustive list biological processes is presented in Table 5–10. Pan-expression of these

markers across 7–18 cancers other than colorectal in origin (Table 5–10), however, could prompt false-positive findings. Meanwhile, CA125 is used in clinical practice to differentiate ovarian cancers and its use for ovarian cancer screening is currently being assessed [55]. It is also elevated during numerous malignancies and benign inflammatory conditions [56,57]. It is unclear why CA125 positivity was greater during preclinical benign adenoma (10–20%) than for CRC (5–8%). Perhaps inflammation is required only during the incipient stages of carcinogenesis, or that malignancies gain the ability to evade detection from immunity.

Table 5–10: Biological processes of shortlisted proteins.

Protein	Verified biological processes[1]	Positivity in other malignancies[2] (tissue) [58]
CA125	Cell adhesion	Breast ; Cervical; Endometrial; Gastric; Ovarian; Pancreatic; Thyroid; Urothelial;
CEA	Cell adhesion; Leukocyte migration; Negative regulation of anoikis/apoptosis/myotube differentiation	Breast; Cervical; Liver; Lung; Ovarian; Pancreatic; Stomach;
CYFRA21-1	[For Cytokeratin 19] Response to oestrogen; Sacromere organisation	Breast; Carcinoid; Cervical; Endometrial; Glioma; Head & neck; Liver; Lung; Melanoma; Ovarian; Pancreatic; Prostate; Renal; Skin; Stomach; Testis; Thyroid; Urothelial
FAP	Endothelial cell migration; Melanocyte apoptosis/proliferation; Negative regulation of cell proliferation involved in contact inhibition/extracellular matrix disassembly/extracellular matrix organisation; Positive regulation of cell-cycle arrest/execution-phase apoptosis; Proteolysis; Regulation of collagen catabolic process/fibrinolysis	Breast; Carcinoid; Cervical; Endometrial; Glioma; Head & neck; Liver; Lung; Melanoma; Ovarian; Pancreatic; Prostate; Renal; Skin; Stomach; Testis; Thyroid; Urothelial
TIMP-1	Negative regulation of catalytic activity/endopeptidase activity/membrane ectodomain proteolysis/metalloproteinase activity; Positive regulation of cell population proliferation; Regulation of integrin-mediated signalling pathway	Breast; Carcinoid; Cervical; Endometrial; Glioma; Liver; Lung; Lymphoma; Ovarian; Pancreatic; Prostate; Stomach; Testis
VEGFA	Angiogenesis; Negative regulation of apoptotic process; Positive regulation of cell adhesion/endothelial cell migration/endothelial cell proliferation/focal adhesion assembly; Response to hypoxia; Vascular endothelial growth factor signalling pathway; Vasculogenesis	Breast; Carcinoid; Cervical; Endometrial; Glioma; Head & Neck; Liver; Lung; Lymphoma; Melanoma; Ovarian; Pancreatic; Prostate; Renal; Skin; Stomach; Testis; Thyroid; Urothelial

Annotations listed by [1] https://www.uniprot.org and [2] https://v15.proteinatlas.org/

Implications for clinicians and policymakers. The often-criticised gFOBT is estimated to be 50% sensitive for preclinical CRCs diagnosed within a year and 25% within four years of the test [59]. Nonetheless, this compares favourably to the performance of CEA, which is demonstrated here to be 25% and 12% sensitive at the same lead-time intervals. Moreover, since concluding this study, the UK National Screening Committee has recommended that the BCSP should replace the gFOBT with the next-generation FIT [60]. Studies have demonstrated thus far that within the BCSP the FIT increased overall uptake by 7% overall and by 11% in previous non-responders [9]. It also has superior sensitivity over the gFOBT [61]. Thus, it seems unlikely for CEA to be suitable for CRC screening unless its sensitivity can be supplemented.

Unanswered questions and future research. A nuance of applying a population-based static threshold is that it often overlooks individual variation in marker levels. It has thus been proposed that deviance from an individualised baseline suggestive of exponential tumour growth could be cause for investigation [62]. The preliminary analysis presented here suggests that the temporal behaviour of CEA may warrant further attention based on its high variability in preclinical Dukes C–D cancers versus women who remained cancer-free or who were later diagnosed with benign adenoma or Dukes A–B CRC. The ROCA (Risk of Ovarian Cancer Algorithm) for screening for ovarian cancer, for example, is markedly superior to screening based on a static CA125 threshold [63]. The algorithm takes a Bayesian approach to estimate the personal risk of ovarian cancer based on risks derived from age-specific incidences and a 'changepoint' in CA125 over time from annual measurements [62]. The same model could be fit to preclinical CEA measurements. Doing so would present a number of logistical issues: would annual measurements (over current biennial testing) be exorbitant to cost effectiveness? Or would any improvement in sensitivity be justified?

Failures in developing markers are numerous [64,65]. Indeed, the present study highlights the need to select potential candidates based on their performance in preclinical samples, with otherwise promising markers failing to live up to their diagnostic performances. The screening sensitivities of CYFRA21-1, FAP, TIMP-1, and VEGFA (12%, 7%, 12% and 12%, respectively) at 0–1 years before diagnosis, for example, were poor in comparison to their sensitivities for detecting Dukes A–B cancers at diagnosis (37%, 44%, 42–48%, and 54%) [19,20,23,24,41]. Future studies aiming to identify novel screening markers for CRC should, therefore, use preclinical samples taken from

asymptomatic patients in advance of their diagnosis in order to increase the odds of success.

5.5 Conclusions

CEA is elevated in a subset of patients with preclinical CRC but is, by itself, not a viable screening test. No existing markers have the required performance to be used individually for screening for CRC in asymptomatic patients. Future studies aiming to discover novel screening markers should use preclinical samples to enhance the odds of success.

5.6 References

[1] D.S. Thomas, E.-O. Fourkala, S. Apostolidou, R. Gunu, A. Ryan, I. Jacobs, U. Menon, W. Alderton, A. Gentry-Maharaj, J.F. Timms, Evaluation of serum CEA, CYFRA21-1 and CA125 for the early detection of colorectal cancer using longitudinal preclinical samples, Br. J. Cancer. 113 (2015) 268–74. doi:10.1038/bjc.2015.202.

[2] M. Navarro, A. Nicolas, A. Ferrandez, A. Lanas, Colorectal cancer population screening programs worldwide in 2016: an update, World J. Gastroenterol. 23 (2017) 3632–3642. doi:10.3748/wjg.v23.i20.3632.

[3] E.H. Schreuders, A. Ruco, L. Rabeneck, R.E. Schoen, J.J.Y. Sung, G.P. Young, E.J. Kuipers, Colorectal cancer screening: a global overview of existing programmes, Gut. 64 (2015) 1637–1649. doi:10.1136/gutjnl-2014-309086.

[4] P. Hewitson, P. Glasziou, E. Watson, B. Towler, L. Irwig, Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update, Am J Gastroenterol. 103 (2008) 1541–1549. doi:10.1111/j.1572-0241.2008.01875.x.

[5] C. von Wagner, G. Baio, R. Raine, J. Snowball, S. Morris, W. Atkin, A. Obichere, G. Handley, R.F. Logan, S. Rainbow, S. Smith, S. Halloran, J. Wardle, Inequalities in participation in an organized national colorectal cancer screening programme: results from the first 2.6 million invitations in England, Int. J. Epidemiol. 40 (2011) 712–8. doi:10.1093/ije/dyr008.

[6] Y. Hirst, S. Stoffel, G. Baio, L. McGregor, C. von Wagner, Uptake of the English Bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out, Eur. J. Cancer. 103 (2018) 267–273. doi:10.1016/j.ejca.2018.07.135.

[7] A. Ghanouni, S.G. Smith, S. Halligan, A. Plumb, D. Boone, G.L. Yao, S. Zhu, R. Lilford, J. Wardle, C. von Wagner, Public preferences for colorectal cancer screening tests: a review of conjoint analysis studies, Expert Rev. Med. Devices. 10 (2013) 489–499. doi:10.1586/17434440.2013.811867.

[8] A. Adler, S. Geiger, A. Keil, H. Bias, P. Schatz, T. DeVos, J. Dhein, M. Zimmermann, R. Tauber, B. Wiedenmann, Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany, BMC Gastroenterol. 14 (2014) 183. doi:10.1186/1471-230X-14-183.

[9] S. Moss, C. Mathews, T.J. Day, S. Smith, H.E. Seaman, J. Snowball, S.P. Halloran, Increased uptake and improved outcomes of bowel cancer screening with a faecal

immunochemical test: results from a pilot study within the national screening programme in England, Gut. 66 (2017) 1631–1644. doi:10.1136/gutjnl-2015-310691.

[10] C.K. Palmer, M.C. Thomas, C. von Wagner, R. Raine, Reasons for non-uptake and subsequent participation in the NHS Bowel Cancer Screening Programme: a qualitative study, Br. J. Cancer. 110 (2014) 1705–11. doi:10.1038/bjc.2014.125.

[11] S.H. Lo, J. Waller, J. Wardle, C. von Wagner, Comparing barriers to colorectal cancer screening with barriers to breast and cervical screening: a population-based survey of screening-age women in Great Britain, J. Med. Screen. 20 (2013) 73–79. doi:10.1177/0969141313492508.

[12] E.G. Liles, G.D. Coronado, N. Perrin, A.H. Harte, R. Nungesser, N. Quigley, N.T. Potter, G. Weiss, T. Koenig, T. DeVos, Uptake of a colorectal cancer screening blood test is higher than of a fecal test offered in clinic: a randomized trial, Cancer Treat. Res. Commun. 10 (2017) 27–31. doi:10.1016/J.CTARC.2016.12.004.

[13] M.J. Duffy, R. Lamerz, C. Haglund, A. Nicolini, M. Kalousová, L. Holubec, C. Sturgeon, Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update, Int. J. Cancer. 134 (2014) 2513–2522. doi:10.1002/ijc.28384.

[14] G.J. van Kamp, S. von Mensdorff-Pouilly, P. Kenemans, R. Verstraeten, C.A. Yedema, T. Wobbes, G.G. Bon, A. Kok, J.P. Brown, J. Hilgers, Evaluation of colorectal cancer-associated mucin CA M43 assay in serum, Clin. Chem. 39 (1993) 1029–32.

[15] H.J. Kim, H.J. Kang, H. Lee, S.T. Lee, M.H. Yu, H. Kim, C. Lee, Identification of S100A8 and S100A9 as serological markers for colorectal cancer, J Proteome Res. 8 (2009) 1368–1379. doi:10.1021/pr8007573.

[16] R. Palmqvist, B. Engarås, G. Lindmark, G. Hallmans, B. Tavelin, O. Nilsson, S. Hammarström, L. Hafström, Prediagnostic levels of carcinoembryonic antigen and CA 242 in colorectal cancer: a matched case-control study, Dis. Colon Rectum. 46 (2003) 1538–44. doi:10.1097/01.DCR.0000089121.59268.D2.

[17] L.C. Fernandes, S.B. Kim, D. Matos, Cytokeratins and carcinoembryonic antigen in diagnosis, staging and prognosis of colorectal adenocarcinoma, World J Gastroenterol. 11 (2005) 645–648.

[18] N.J. Shimwell, W. Wei, S. Wilson, M.J. Wakelam, T. Ismail, T. Iqbal, P.J. Johnson, A. Martin, D.G. Ward, Assessment of novel combinations of biomarkers for the detection of colorectal cancer, Cancer Biomark. 7 (2010) 123–132. doi:10.3233/cbm-2010-0155.

[19] N. Wild, H. Andres, W. Rollinger, F. Krause, P. Dilba, M. Tacke, J. Karl, A combination of serum markers for the early detection of colorectal cancer, Clin Cancer Res. 16 (2010) 6111–6121. doi:10.1158/1078-0432.CCR-10-0119.

[20] E.T. Waas, T. Hendriks, R.M.L.M. Lomme, T. Wobbes, Plasma levels of matrix metalloproteinase-2 and tissue inhibitor of metalloproteinase-1 correlate with disease stage and survival in colorectal cancer patients, Dis. Colon Rectum. 48 (2005) 700–10.

[21] B. Knychalski, T. Lukieńczuk, The evaluation of diagnostic value of the tumor markers: CCSA-2 and CEA in colorectal cancer, Pol. Przegl. Chir. 84 (2012) 86–92. doi:10.2478/v10035-012-0014-3.

[22] M. Roessler, W. Rollinger, S. Palme, M.-L. Hagmann, P. Berndt, A.M. Engel, B. Schneidinger, M. Pfeffer, H. Andres, J. Karl, H. Bodenmüller, J. Rüschoff, T. Henkel, G. Rohr, S. Rossol, W. Rösch, H. Langen, W. Zolg, M. Tacke, Identification of nicotinamide

N-methyltransferase as a novel serum tumor marker for colorectal cancer, Clin. Cancer Res. 11 (2005) 6550–7. doi:10.1158/1078-0432.CCR-05-0983.

[23] M.N. Holten-Andersen, I.J. Christensen, H.J. Nielsen, R.W. Stephens, V. Jensen, O.H. Nielsen, S. Sørensen, J. Overgaard, H. Lilja, A. Harris, G. Murphy, N. Brünner, Total levels of tissue inhibitor of metalloproteinases 1 in plasma yield high diagnostic sensitivity and specificity in patients with colon cancer, Clin. Cancer Res. 8 (2002) 156–64.

[24] J.H. Lee, Clinical Usefulness of Serum CYFRA 21–1 in Patients with Colorectal Cancer, Nucl. Med. Mol. Imaging (2010). 47 (2013) 181–187. doi:10.1007/s13139-013-0218-4.

[25] J. Schneider, N. Bitterlich, G. Schulze, Improved sensitivity in the diagnosis of gastro-intestinal tumors by fuzzy logic-based tumor marker profiles including the tumor M2-PK, Anticancer Res. 25 (2005) 1507–15.

[26] J.J. Ladd, T. Busald, M.M. Johnson, Q. Zhang, S.J. Pitteri, H. Wang, D.E. Brenner, P.D. Lampe, R. Kucherlapati, Z. Feng, R.L. Prentice, S.M. Hanash, Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women, Cancer Prev. Res. (Phila). 5 (2012) 655–64. doi:10.1158/1940-6207.CAPR-11-0412.

[27] M.S. Pepe, R. Etzioni, Z. Feng, J.D. Potter, M.L. Thompson, M. Thornquist, M. Winget, Y. Yasui, Phases of biomarker development for early detection of cancer, J. Natl. Cancer Inst. 93 (2001) 1054–61.

[28] M.S. Pepe, Z. Feng, H. Janes, P.M. Bossuyt, J.D. Potter, Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design, J. Natl. Cancer Inst. 100 (2008) 1432–8. doi:10.1093/jnci/djn326.

[29] U. Menon, A. Gentry-Maharaj, A. Ryan, A. Sharma, M. Burnell, R. Hallett, S. Lewis, A. Lopez, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, S. Skates, M. Parmar, I. Jacobs, Recruitment to multicentre trials—lessons from UKCTOCS: descriptive study, BMJ. 337 (2008) a2079–a2079. doi:10.1136/bmj.a2079.

[30] R Core Team, R: a language and environment for statistical computing, (2016). https://www.r-project.org/.

[31] C. Ritz, F. Baty, J.C. Streibig, D. Gerhard, Dose-Response Analysis Using R, PLoS One. 10 (2015) e0146021. doi:10.1371/journal.pone.0146021.

[32] M. Stevenson, T. Nunes, C. Heuer, J. Marshall, J. Sanchez, R. Thornton, J. Reiczigel, J. Robison-Cox, P. Sebastiani, P. Solymos, K. Yoshida, G. Jones, S. Pirikahu, S. Firestone, R. Kyle, J. Popp, M. Jay, epiR: tools for the Analysis of Epidemiological Data, (2017). https://cran.r-project.org/package=epiR.

[33] H. Wickham, ggplot2: elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.

[34] F.E. Harrell, Hmisc: Harrell Miscellaneous, (2018). https://cran.r-project.org/package=Hmisc.

[35] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinformatics. 12 (2011) 77. doi:10.1186/1471-2105-12-77.

[36] M. López-Ratón, M.X. Rodríguez-Álvarez, C.C. Suárez, F.G. Sampedro, OptimalCutpoints: an R Package for Selecting Optimal Cutpoints in Diagnostic Tests, J. Stat. Softw. 61 (2014) 1–36. doi:10.18637/jss.v061.i08.

[37] H. Wickham, L. Henry, tidyr: easily Tidy Data with "spread()" and "gather()" Functions, (2018). https://cran.r-project.org/package=tidyr.

[38] M. Groblewska, B. Mroczko, M. Gryko, B. Kędra, M. Szmitkowski, Matrix metalloproteinase 2 and tissue inhibitor of matrix metalloproteinases 2 in the diagnosis of colorectal adenoma and cancer patients, Folia Histochem. Cytobiol. 48 (2010) 564–71. doi:10.2478/v10042-010-0076-1.

[39] A. Suppiah, A. Alabi, L. Madden, J.E. Hartley, J.R.T. Monson, J. Greenman, Anti-p53 autoantibody in colorectal cancer: prognostic significance in long-term follow-up, Int. J. Colorectal Dis. 23 (2008) 595–600. doi:10.1007/s00384-008-0458-4.

[40] S. Tao, U. Haug, K. Kuhn, H. Brenner, Comparison and combination of blood-based inflammatory markers with faecal occult blood tests for non-invasive colorectal cancer screening, Br. J. Cancer. 106 (2012) 1424–30. doi:10.1038/bjc.2012.104.

[41] W.-S. Tsai, C.R. Changchien, C.-Y. Yeh, J.-S. Chen, R. Tang, J.-M. Chiang, P.-S. Hsieh, C.W. Fan, J.Y. Wang, Preoperative plasma vascular endothelial growth factor but not nitrite is a useful complementary tumor marker in patients with colorectal cancer, Dis. Colon Rectum. 49 (2006) 883–94. doi:10.1007/s10350-006-0528-z.

[42] N. Duraker, D. Can, M. Parilti, Measurement of serum total and free prostate-specific antigen in women with colorectal carcinoma, Br. J. Cancer. 86 (2002) 203–6. doi:10.1038/sj.bjc.6600049.

[43] B. Knychalski, T. Lukienczuk, The evaluation of diagnostic value of the tumor markers: CCSA-2 and CEA in colorectal cancer, Pol Przegl Chir. 84 (2012) 86–92. doi:10.2478/v10035-012-0014-3.

[44] P. Kuusela, C. Haglund, P.J. Roberts, Comparison of a new tumour marker CA 242 with CA 19-9, CA 50 and carcinoembryonic antigen (CEA) in digestive tract diseases, Br. J. Cancer. 63 (1991) 636–40.

[45] X. Filella, A.M. Ballesta, M. Fox, H. Mitchell, R. Molina, P. Pürstner, H. Thome, Multicentre clinical evaluation of the COBAS CORE CEA, CA 125 II and PSA tumor marker assays, Int. J. Biol. Markers. 11 (1996) 40–5.

[46] Cancer Research UK, Bowel cancer incidence statistics [UK, 2015, ICD-10 C18–C20], (2018). http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence (accessed February 28, 2018).

[47] H. Shinya, W.I. Wolff, Morphology, anatomic distribution and cancer potential of colonic polyps, Ann. Surg. 190 (1979) 679–83.

[48] C. Kut, F. Mac Gabhann, A.S. Popel, Where is VEGF in the body? A meta-analysis of VEGF distribution in cancer, Br. J. Cancer. 97 (2007) 978–85. doi:10.1038/sj.bjc.6603923.

[49] K. Jung, L. Nowak, M. Lein, W. Henke, D. Schnorr, S.A. Loening, Role of specimen collection in preanalytical variation of metalloproteinases and their inhibitors in blood, Clin. Chem. 42 (1996) 2043–5.

[50] M.J. Duffy, Carcinoembryonic antigen as a marker for colorectal cancer: is it clinically useful?, Clin. Chem. 47 (2001) 624–30. http://www.ncbi.nlm.nih.gov/pubmed/11274010 (accessed March 10, 2013).

[51] E. Puré, R. Blomberg, Pro-tumorigenic roles of fibroblast activation protein in cancer: back to the basics, Oncogene. 37 (2018) 4343–4357. doi:10.1038/s41388-018-0275-3.

[52] K. Kessenbrock, V. Plaks, Z. Werb, Matrix Metalloproteinases: regulators of the Tumor Microenvironment, Cell. 141 (2010) 52–67. doi:10.1016/j.cell.2010.03.015.

[53] D. Hanahan, R.A. Weinberg, The hallmarks of cancer, Cell. 100 (2000) 57–70.

[54] V. Barak, H. Goike, K.W. Panaretakis, R. Einarsson, Clinical utility of cytokeratins as tumor markers, Clin. Biochem. 37 (2004) 529–540. doi:10.1016/j.clinbiochem.2004.05.009.

[55] I.J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J.K. Kalsi, N.N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, D.N. Crump, S.K. Davies, A. Dawnay, S. Dobbs, G. Fletcher, J. Ford, K. Godfrey, R. Gunu, M. Habib, R. Hallett, J. Herod, H. Jenkins, C. Karpinskyj, S. Leeson, S.J. Lewis, W.R. Liston, A. Lopes, T. Mould, J. Murdoch, D. Oram, D.J. Rabideau, K. Reynolds, I. Scott, M.W. Seif, A. Sharma, N. Singh, J. Taylor, F. Warburton, M. Widschwendter, K. Williamson, R. Woolas, L. Fallowfield, A.J. McGuire, S. Campbell, M. Parmar, S.J. Skates, Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial, Lancet. 387 (2016) 945–956. doi:10.1016/S0140-6736(15)01224-6.

[56] H. Ataseven, Z.A. Öztürk, M. Arhan, O. Yüksel, S. Köklü, M. Ibiş, Ö. Başar, F.M. Yılmaz, I. Yüksel, Cancer antigen 125 levels in inflammatory bowel diseases, J. Clin. Lab. Anal. 23 (2009) 244–248. doi:10.1002/jcla.20323.

[57] I. Jacobs, R.C. Bast, The CA 125 tumour-associated antigen: a review of the literature, Hum. Reprod. 4 (1989) 1–12.

[58] M. Uhlen, L. Fagerberg, B.M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjostedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C.A.-K. Szigyarto, J. Odeberg, D. Djureinovic, J.O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J.M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Ponten, Tissue-based map of the human proteome, Science (80-.). 347 (2015) 1260419–1260419. doi:10.1126/science.1260419.

[59] J.E. Allison, R. Feldman, I.S. Tekawa, Hemoccult screening in detecting colorectal neoplasm: sensitivity, specificity, and predictive value. Long-term follow-up in a large group practice setting, Ann. Intern. Med. 112 (1990) 328–33.

[60] UK National Screening Committee, Note of the meeting held on the 19 November 2015 at Goodenough College, London, 2015.

[61] L. Hol, J.A. Wilschut, M. van Ballegooijen, A.J. van Vuuren, H. van der Valk, J.C.I.Y. Reijerink, A.C.M. van der Togt, E.J. Kuipers, J.D.F. Habbema, M.E. van Leerdam, Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels, Br. J. Cancer. 100 (2009) 1103–1110. doi:10.1038/sj.bjc.6604961.

[62] S.J. Skates, D.K. Pauler, I.J. Jacobs, Screening Based on the Risk of Cancer Calculation From Bayesian Hierarchical Changepoint and Mixture Models of Longitudinal Markers, J. Am. Stat. Assoc. 96 (2001) 429–439. doi:10.1198/016214501753168145.

[63] U. Menon, A. Ryan, J. Kalsi, A. Gentry-Maharaj, A. Dawnay, M. Habib, S. Apostolidou, N. Singh, E. Benjamin, M. Burnell, S. Davies, A. Sharma, R. Gunu, K. Godfrey, A. Lopes, D. Oram, J. Herod, K. Williamson, M.W. Seif, H. Jenkins, T. Mould, R. Woolas, J.B. Murdoch, S. Dobbs, N.N. Amso, S. Leeson, D. Cruickshank, I. Scott, L. Fallowfield, M. Widschwendter, K. Reynolds, A. McGuire, S. Campbell, M. Parmar, S.J. Skates, I. Jacobs, Risk Algorithm Using Serial Biomarker Measurements Doubles the Number of Screen-Detected Cancers Compared With a Single-Threshold Rule in the United Kingdom Collaborative Trial of Ovarian Cancer Screening, J. Clin. Oncol. 33 (2015) 2062–2071. doi:10.1200/JCO.2014.59.4945.

[64] S.E. Kern, Why Your New Cancer Biomarker May Never Work: Recurrent Patterns and Remarkable Diversity in Biomarker Failures, Cancer Res. 72 (2012) 6097–6101. doi:10.1158/0008-5472.CAN-12-3232.

[65] E.P. Diamandis, The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem?, BMC Med. 10 (2012) 87. doi:10.1186/1741-7015-10-87.

Chapter 6. Identifying Novel Markers for The Early Detection of Colorectal Cancer Through LC-MS/MS Analysis of Serial Preclinical Sera

A Retrospective Case-Control Study Nested Within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

6.1 Background

CEA is the most promising of many circulating markers that, individually, have no use for screening asymptomatic individuals for CRC. If, however, blood-based screening is to be a viable alternative to the gFOBT or FIT, novel markers are needed to complement, or even surpass, the performance of CEA. Nonetheless, novel markers are likely to be 'high-hanging fruit'; present in trace amounts having been secreted or leaked into the large volume of the circulatory system by a multiplying, asymptomatic tumour.

Numerous prospective studies have so-far procured the preferred preclinical samples, and nested case-control studies have since reported the application of microarrays and mass spectrometry to identify differentially expressed markers from these preclinical sample sets (Table 6–1). A number of markers have been shortlisted, and a combination of these are being taken forward as part of multi-analyte panels.

A longstanding challenge for sampling the low-abundance proteome is to circumvent the Pareto distribution of protein abundance in serum/plasma wherein the 22 most abundant account for 99% of the total mass and are an estimated 12 orders of magnitude more abundant than the thousands of low-abundance proteins [5,6]. Samples have thus been enriched and simplified via the removal of high-abundance proteins [7] and fractionated downstream according to analyte polarity/hydrophobicity and charge [8].

The present study aimed to identify markers that supersede or complement the performance of existing markers. It reports on the analysis of pooled preclinical serial sera though the application of immunodepletion, TMTs, and three-dimensional LC-MS/MS.

Table 6–1: An overview of current studies using preclinical samples to discover markers for CRC screening.

Serum repository	Study set	Methodology
Women's Health Initiative (WHI) [1]	*Subjects.* Women aged 65 years on average enrolled to the WHI. *Samples.* 90 plasma drawn ≤ 18 months before diagnosis and 90 drawn from matched controls with no cancer history.	LC-MS/MS analyses (LTQ-Orbitrap XL) with acrylamide isotope labelling.
European Prospective Investigation into Cancer and Nutrition (EPIC) [2]	*Subjects.* Male and females aged 53 on average enrolled to the EPIC centre in Florence, Italy. *Samples.* Ten preclinical plasma samples (lead-time undefined) and ten plasma from matched controls who remained cancer-free at time of case diagnosis.	LC-MS/MS analyses (LTQ Orbitrap XL) with label-free quantitation.
USA Department of Defence Serum Repository (DoDSR) [3]	*Subjects.* Male and female US military personal aged 39.5 years on average who remained on active duty. *Samples.* 2,752 serial sera (3.5 sampler per person on average).	MALDI-TOF/TOF with label-free quantitation.
Cardiovascular Health study [4]	*Subjects.* Males and females aged ≥ 65 years, excluding those diagnosed with myocardial infarction, angina pectoris or stroke since enrolment. *Samples.* 79 plasmas drawn ≤ 36 months before diagnosis and 79 from matched controls who remained cancer-free.	High-density array of ~3,200 antibodies with affinity for > 2,100 proteins.

Abbreviations: LC-MS/MS, liquid chromatography-tandem mass spectrometry; MALDI-TOF-TOF, Matrix-assisted laser desorption/ionisation tandem Time-of-Flight Mass spectrometry).

6.2 Prelude to Methods

The brief prelude that follows, while grossly simplified, is intended to provide sufficient context to make sense of the Mass Spectrometry presented in the remaining chapter. Interested readers are directed to current reviews of mass spectrometry-based proteomics [9,10].

The study of Mass Spectrometry concerns the measurement of mass—and more specifically the atomic mass of gaseous ions. Take a positively charged glycine residue ($C_2H_3NO^+$) as an example. Made of two carbons with a mass of 12.011 Da (Dalton, equal to 1/12 of the mass of a carbon-12 isotope by definition), three hydrogens (1.008 Da), a nitrogen (14.007 Da), an oxygen (15.999 Da), and a proton (1.007 Da), a mass analyser would detect it as an entity with a mass of 58.058 Da. For a peptide of three amino acids, however, mass analyses would yield only the total mass and thus a peptide, once mass analysed, must be fragmented via collision with an inert gas to unveil the individual amino acid components in tandem mass analyses (MS/MS). (Proteins are almost always proteolysed first into peptides amenable to mass analyses using trypsin). An initial mass spectrum of an ionised SVL peptide (Serine–Threonine–Leucine), ignoring for the time being protein groups other than the residual, would show with a mass of 300.192 Da (S 87.032 + V 99.068 + L 113.084 + $^+$ 1.007) and as fragments of SV, VL, S, V, and L in tandem mass spectra. Peptides fragment somewhat predictably at the peptide bond (C–N) linking the amino and carboxyl terminal groups to give b- or y-type ions dependent on the proton charge being retained on the amino- or carboxy-terminal fragments, respectively [11]. Proteins are thereafter identified with a probabilistic algorithm that compares the observed b and y ions with those predicted given a database of human protein sequences proteolysed and fragmented *in silico* [12].

Many proteomic experiments require not only a list of proteins present in a biological sample but also a quantitative measure of their abundance. The relatively recent advent of chemical Tandem Mass Tags provides a high-throughput solution for multiplex quantitation for six or ten experiential conditions [13,14]. Each isobaric reagent is of equal mass to one another but variably consists of a unique mass reporter (in the range of 126–131 Da inclusively for the sixplex variant), a mass normaliser that offsets for the mass differences of the reporter, and an amine-reactive group that enables covalent labelling of lysines and the amino termini of proteins proteolysed with trypsin. Thus, while each experimental arm is labelled with a unique tag, they are combined and submitted in the same MS experiment, where they co-migrate through the mass analysers and on

fragmentation (with an energy higher than required for peptide sequence information) the relative abundances of the mass reporter ions are unveiled. Absolute quantitation, when necessary, can be achieved with a labelled internal standard.

6.3 Methods

Study design. The present population-based, retrospective case-control study was nested within the multimodal arm of the UKCTOCS involving 50,640 women who donated blood annually. Subjects were recruited to the UKCTOCS from NHS patient registers to one of thirteen regional UKCTOCS centres in England, Northern Ireland, and Wales during 2001–05. Women with previous bilateral oophorectomy or ovarian malignancy, heightened risk of familial ovarian cancer, or an active non-ovarian malignancy were excluded from recruitment to the UKCTOCS [15]. Only those who consented to their involvement in secondary studies were eligible for the study.

Subjects. One-hundred subjects were studied, including 40 cases, 20 matched benigns, and 40 matched controls. These were the same cohort studied in the previous chapter where full details are outlined in Chapter 5. Briefly, cases were clinically verified adenocarcinomas of the colon or rectum, excluding those of the appendix and those with a signet-ring cell or mucinous morphology. Benigns were women with a benign diagnosis (D12, excluding those of the appendix (D12.1), anus and anal canal (D12.9)) recorded in their HES and no cancer diagnoses of any origin according to cancer and death registries, HES, or self-reporting. Controls were women who had no diagnoses of cancer of any origin on according to cancer and death registries, HES, and self-reporting as of 1 September 2013. The baseline characteristics of subjects and clinical and histological features of benign adenomas and adenocarcinomas were presented previously (Table 5–5 & Table 5–6, respectively).

Samples. The discovery experiment used a set of 200 serial preclinical sera (80 controls, 40 benign, 80 cancers) drawn from 100 subjects (of whom 40 developed CRC, 20 benign, and 40 controls who were free of cancer at study commencement). Samples predated the diagnosis of benign adenoma or CRC by 0–1 and 3–4 years (Table 6–2).

Table 6–2: Sample counts according to lead time and clinical group.				
Years lead time	**Control**	**Benign**	**Dukes A–B**	**Dukes C–D**
0–1	40	20	20	20
3–4	40	20	20	20
Total	80	40	40	40

Marker candidates were verified on 386 serial preclinical sera (154 controls, 78 benign, 154 cancers) drawn from 100 subjects (of whom 40 developed CRC, 20 benign, and 40 controls who were free of cancer at study commencement). Samples predated the diagnosis of benign adenoma or CRC by 0–4 years inclusively. These were drawn from the same 100 subjects included in the discovery experiment and were the same sample set used in Chapter 5.

Bloods were drawn from UKCTOCS multimodal women by venepuncture, separated to serum according to a standardised protocol, and then transported frozen in heat-sealed straws at the end of each day to a cryorespository for long-term storage at –196°C [16]. For this study, samples (i.e. one straw per sample) were retrieved and shipped to the laboratory on dry ice, where they were pseudonymised, thawed at 4°C, randomly aliquoted into 96-well plates, and stored at –80°C until processed for study.

Marker discovery

Pooling. Sera were thawed gently and combined in equal volumes to their respective pool (each pool consisting of 20 sera). These were: 1) 0–1 year preceding Dukes A–B cancers 2) 3–4 years preceding Dukes A–B cancers; 3) 0–1 year for controls matched to Dukes A–B cancers; 4) 3–4 years for controls matched to Dukes A–B cancers; 5) 0–1 year preceding benign adenoma; 6) 3–4 years preceding benign adenoma; 7) 0–1 year preceding Dukes C–D cancers 8) 3–4 years preceding Dukes C–D cancers; 9) 0–1 year for controls matched to Dukes C–D cancers; and 10) 3–4 years for controls matched to Dukes C–D cancers. Notably, sera drawn from benigns were analysed in both sequences and would provide an internal reference for inter-sequence comparisons. The protein content of each of pool was determined using the bicinchoninic acid assay using bovine serum albumin as the standard (Thermo Fisher Scientific Ltd.)

Immunodepletion. The 20 high-abundance proteins that account for ~97% serum protein were removed from each serum pool by antibody-based depletion using two commercial kits completed according to the manufacturer's instructions. Briefly, each 1 mg serum protein was incubated with 1 mL of Proteome Purify 12 resin slurry (R&D

Systems) for 30 min. The eluent was recovered and then concentrated to 100 µL by membrane ultrafiltration at 4°C using a 5-kDa molecular weight cut-off Vivaspin column (GE Healthcare). Next, the concentrate was incubated with an equilibrated ProteoPrep 20 spin column (Sigma-Aldrich) for 20 min. The eluent was recovered by centrifugation and then vacuum-dried. The resulting pellet was re-suspended in 100 mM tetraethylammonium bicarbonate (TEAB; pH 8.5) containing 0.1% SDS. The protein content was assayed using the bicinchoninic acid assay with bovine serum albumin standards. The efficiency of immunodepletion was calculated as 1 – (concentration post immunodepletion/initial concentration).

Proteolytic digestion. Proteins were digested into peptides using trypsin. For this, 73 µg of the recovered protein was suspended in 100 µL of 100 mM TEAB (pH 8.5) containing 0.1% SDS, reduced with 1 mM tris(2-carboxyethyl)phosphine over 1 h at 55°C, alkylated with 7.5 mM 2-iodoacetamide for 1 h at room temperature in darkness, and digested with 1 part porcine trypsin (Promega) 25 parts protein overnight at 37°C.

Tandem Mass Tag labelling. Technological limitations constrained to the use of six unique isobaric tags per run maximum—necessitating two sequence runs for the ten pooled groups (benign 0–1 and 3–4 were run in each sequence to allow between sequence comparisons). Tandem mass tags (TMTs; Thermo Scientific) sixplex 126–131 Da were equilibrated to room temperature and suspended in 41 µL acetonitrile (17.5 mM) and then incubated with the assigned peptide pool for 1 h at room temperature. Next, the labelling reaction was quenched at room temperature over 30 min with hydroxylamine (0.25% final concentration). The TMT-labelled samples were then collected by centrifugation and combined in equal amounts to their assigned sequence pool, each with six TMTs. Excess detergent and salts were removed from the now-pooled labelled peptides using 0.5 mL Pierce Detergent Removal resin (Thermo Scientific) and 1-cc, 30-µm particle size Oasis HLB solid-phase extraction (Waters) columns sequentially as per the manufacturer's instructions. Labelled peptides were vacuum-dried and re-suspended in 1 mL of 2% phosphoric acid twice over.

Strong Anion Exchange. The overall peptide load was fractionated according to net charge using a strong anion exchange (SAX) resin. Briefly, 219 µL (50 µL resin per 100 µg peptide) Ceramic HyperD F DEAE sorbent (Pall Life Sciences) was activated with

300 µL 1 M NaCl in 20 mM TEAB (pH 8.5) and then drained by centrifugation at 600 g for 1 min. The resin was then washed/centrifuged three times with 300 µL 200 mM TEAB (pH 8.5) and finally equilibrated/centrifuged with 300 µL 20 mM TEAB (pH 8.5). Labelled peptides were incubated with the SAX resin for 5 min with gentle agitation before being collected by centrifugation at 600 g for 1 min. The incubation/centrifugation step was repeated once more. Bound peptides were then step-eluted with 300 µL TEAB (pH 8.5) buffer containing either 50, 100, 200 or 800 mM NaCl over 5 min twice over, with each eluent being collected by centrifugation. Each of the five eluents was brought to 2% phosphoric acid and desalted using 1-cc, 30-µm particle size Oasis HLB solid-phase extraction columns as before.

High-performance Liquid Chromatography. Each of the five SAX fractions was further resolved into twenty fractions by high-pH, reverse-phase high-performance liquid chromatography (HPLC) using an Agilent 1100 HPLC system equipped with a PoroShell300 Extend-C18 column (2.1 x 75 mm, 5 µm-bead size, 300 Å-pore size, Agilent). Mobile phases A and B were 20 mM ammonium formate (pH 8.4) and acetonitrile containing 20 mM ammonium formate (pH 8.4), respectively. The gradient was as follows (200 µL min^{-1}): 0–1 min: 3% solvent B; 1–35 min: 45% solvent B; 35–45 min: 85% solvent B; 45–55 min: 3% solvent B. Fractions were vacuum-dried and then washed/dried three times with 1 mL water before being applied to a C18 ZipTip (Millipore) as per the manufacturer's instructions. Finally, peptides were vacuum-dried and suspended in 7 µL 0.1% formic acid in preparation for LC-MS/MS analysis.

Liquid Chromatography-tandem Mass Spectrometry. An Ultimate 3000 LC system (Dionex) was configured upstream of a LTQ Orbitrap XL hybrid mass spectrometer (Thermo Scientific) equipped with a Picoview nanospray source and controlled with Chromeleon Xpress V. 6.8, Thermo Tune Plus, and Thermo Xcalibur software. Mobile phases A and B were 0.1% formic acid and 99.9 acetonitrile/0.1% formic acid, respectively. Peptides (5 µL injection volume) were injected onto an Acclaim PepMap100 C18 pre-column (5 µm, 100 Å, 300 µm x 5 mm, ThermoFisher Scientific) and washed for 3 min with 10% buffer B at a flow rate of 25 µL min^{-1} and then resolved downstream on an Acclaim PepMap100 C18 nano-LC column (3 µm, 100 Å, 75 µm x 250 mm, ThermoFisher Scientific) maintained at 30°C. Bound peptides were eluted with a 90-min linear gradient of 10–50% buffer B at a flow rate of 300 nL min^{-1}). Survey full-scan mass

spectra (from m/z 400–2,000) were acquired in the Orbitrap at a resolution of 60,000 at m/z 400, followed by top 3 Collision-Induced Dissociation (CID)/Higher-energy Dissociation (HCD) data-dependent acquisition using a normalised collision energy of 40% for HCD and a resolution of 7,500 for detection of product ions in the Orbitrap. Dynamic exclusion was enabled with a list size of 500, excluding for 15 sec. Atmospheric polymethylcyclosiloxane was used as a lock mass (455.12003 m/z) for internal calibration.

MS/MS ion search. Mass spectra were submitted via Proteome Discoverer (V 1.4.1.12) to the Mascot search engine for protein identification and quantitation *in silico*, using the SwissProt database (V 2.6, January 2017). Searches were restricted to *Homo sapiens* and were tolerant for ≤ two missed trypsin cleavages, carbamidomethylation of cysteines as a fixed modification, and acetylation of protein N-termini, deamidation of asparagine and glutamine, oxidation of methionine, and TMT-sixplex labelling of lysines and the N-termini of peptides as variable modifications. Precursor and fragment ion searches were tolerant to 10 ppm and 0.5 Da/20 mmu for CID/HCD-spectra, respectively. The Mascot significance and peptide score thresholds were $p < 0.05$ and 20, respectively.

Marker verification

Dermcidin (DCD) and S100A8 were verified on the full set of 386 sera used in the previous chapter to verify CA125, CEA, CYFRA21-1, FAP, TIMP-1, and VEGFA. Briefly, these were 78 and 154 serial sera drawn 0–4 years before the diagnosis of benign adenoma or CRC and 154 serial sera drawn over 0–4 years from controls who remained cancer-free during the study period.

Immunoassays. Serum DCD and S100A8 were determined using commercial immunoassay kits manufactured by Cloud-Clone Corp. (cat no: SEC896Hu) and Thermo Fisher Scientific (cat no: DTM100), respectively. Assays were carried out according to the manufacturer's instructions. Standards ranged from 0.78–50 ng mL^{-1} (1:2 serial dilution) for DCD and 10.24 pg mL^{-1} (1:2.5 serial dilution) for S100A8. Samples were diluted 1:10 for DCD and 1:5 for S100A8. Samples were measured once and standards twice. Absorbance measurements were normalised to that of a background reaction containing only buffer. Determinations were made by interpolation from a four-parametric logistic curve fitted to standards measured on the same plate and corrected for dilution factor. DCD is presented as ng mL^{-1} and S100A8 as pg mL^{-1}.

Data analyses

Inter-sequence correlation. Pools of sera drawn 0–1 and 3–4 years before the diagnosis of benign adenoma acted as technical replicates and were analysed in both LC-MS/MS sequences. Correlation in the relative quantitation (benign 0–1: benign 3–4) between the first and second sequences was assessed using the Spearman's rho after listwise deletion of proteins with missing quantitation in these channels.

Shortlisting of candidate markers. Relative quantitation was expressed as a ratio of four comparisons termed a–d. These were the relative longitudinal differences (i.e. 0–1 vs. 3–4) in cases (a) and controls (b) and the relative difference between cases and controls at 0–1 (c) and 3–4 years (d). Proteins without complete data in all quantitation channels were excluded along with high-abundance proteins that were targeted for immunodepletion (Supplementary Table S1) or that were a subunit of haemoglobin or supposed contaminant (Supplementary Table S2) compiled in the common Repository of Adventitious Proteins (https://www.thegpm.org/cRAP/) or by the Max Plank Institute (http://www.maxquant.org/downloads.htm). Proteins identified with < 2 unique peptides and < 4 peptide spectral matches were deemed to be of low quality and were excluded. Ratios were individually subtracted from 1, transformed using the square-root of square $\sqrt{(x)^2}$, and normalised to a scale of between 0–1 relative to the minimum and maximum values (x – min)/(max – min). Finally, a composite 'Foldscore' was calculated as:

$$\text{Foldscore} = \frac{(xa + xc + xd + (1\text{-}xb))}{4} \quad \textbf{(1)}$$

Foldscore was designed to reward proteins differentially expressed between cases (c) and controls (d) and during cancer growth (a) while penalising those with an unstable baseline in controls (b). Proteins were thereafter ranked in order of descending Foldscore from 1 to 0 (1 being high).

Differentiation. Differences in marker levels between clinical groups were assessed using multivariable analysis (Kruskal-Wallis & post-hoc pairwise Mann-Whitney with

Bonferroni correction). Boxplots denoting the median, IQR, range, and outliers were plotted.

Classification. Cut-off thresholds were calculated at 95% specificity (i.e. a 5% false-positive rate) assuming both up and downregulation during cancer. The sensitivity (TP/(TP+FN)) of each marker were computed with the threshold yielding the highest AUC.

Longitudinal profiles. Average marker serial profiles were modelled with Loess regression. Lead time was estimated by interpolation of Loess curves at the threshold equivalent to 95% specificity (DCD < 0.35 ng mL^{-1}, S100A8 < 164.5 pg mL^{-1})**.** Within-person variation in markers levels across serial samples was assessed using the coefficient of variation (CV). Individual CVs were pooled into controls, benigns, Dukes A–B, and Dukes C–D cancers for which an aggregate median and range were computed. Inter-group differences in individual CVs were assessed using multivariable analysis (Kruskal-Wallis & post-hoc pairwise Wilcox rank sum test with Bonferroni adjustment).

Statistical analyses. All analyses were made with R version 3.3.2 [17] running drc [18], epiR [19], ggplot2 [20], OptimalCutpoints [21], and tidyr [22] packages. A p value < 0.05 was considered statistically significant.

6.4 Results

Marker discovery. Removal of twenty high-abundance proteins reduced the serum protein concentration by 99.9% across all ten pools. Proteins were thereafter digested, labelled with one of six TMTs, combined, fractionated by SAX and high-pH RP-LC, and analysed by LC-MS/MS. The ten pools were analysed in two separate run sequences—hereafter termed 'Early' and 'Late' sequences—with the former consisting of pooled sera drawn from Dukes A–B cancers, matched controls, and benign neoplasms and the latter of pooled sera drawn from Dukes C–D cancers, matched controls, and benign adenomas. There was no evidence of a correlation in the relative quantitation (benign 0–1: benign 3–4) between technical replicates for the same proteins quantitated over the two sequences (rho –0.071; p 0.333). Thus, normalisation to an internal reference would be beneficial for inter-sequence comparisons.

A total of 324 unique protein groups were identified (64 exclusive to the Early sequence, 68 exclusive to the Late sequence, 192 identified in both). Of the 256 proteins identified in the Early sequence, 133 were excluded for not having a complete TMT 126–131 profile (n 12), being a high-abundance protein targeted by immunodepletion (n 27) or a proposed contaminant (n 8), or for being identified with < 2 unique peptides and <4 peptide spectral matches (n 86). Of the 260 proteins identified in the Late sequence, 143 were excluded for not having a complete TMT 126–131 profile (n 28), being a high-abundance protein (n 25) or proposed contaminant (n 11), or for being identified with < 2 unique peptides and < 4 peptide spectral matches (81). After exclusions, 135 unique proteins (18 in the Early sequence, 12 in the Late sequence, 105 in both) were available for consideration.

Proteins were ranked according to their Foldscore, which was an aggregate scoring designed to reward where differential expression is preferential (i.e. in cancers relative to controls or during cancer growth) while penalising those an unstable baseline in controls. Foldscores were normally distributed about a mean Foldscore of 0.28 (Figure 6–1). Five proteins (4 unique) had a Foldscore ≥ 0.4 and were shortlisted for review. These were DCD, alpha-1-acid glycoprotein II (shortlisted in each sequence), carbonic anhydrase I, and S100A8. A selection strategy based on differential expression between cases and controls at 0–1 year only would have selected S100A9 at the expense of carbonic anhydrase I (Table 6–3).

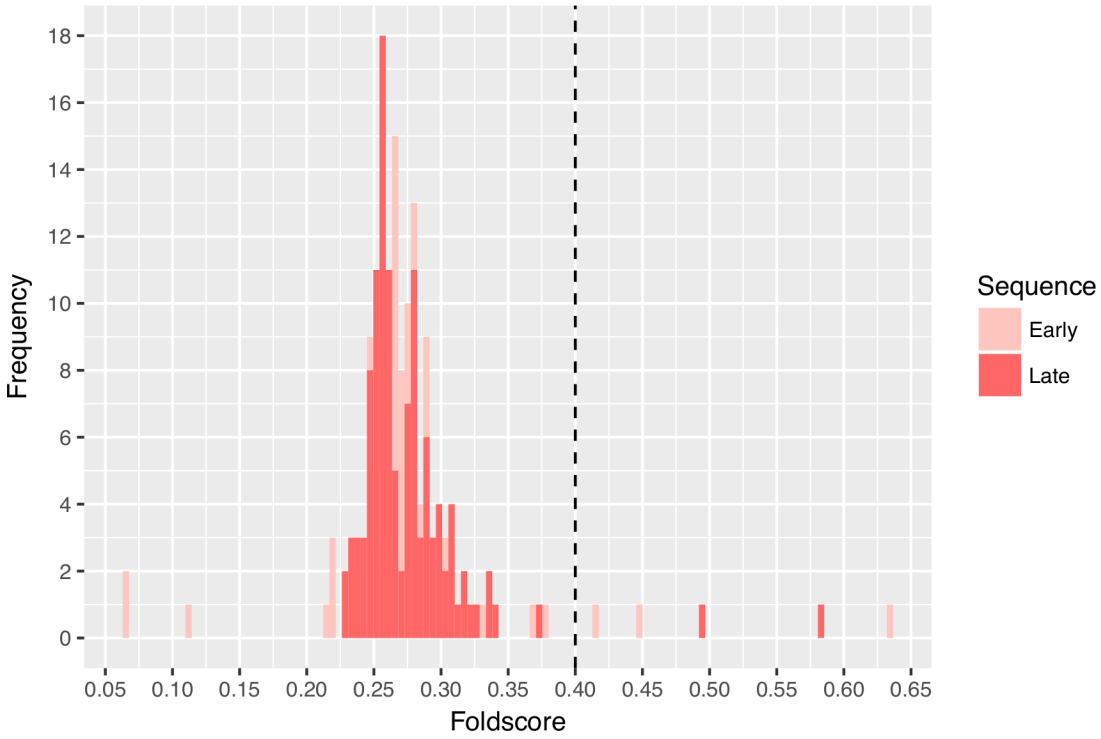


Figure 6–1: Distribution of Foldscores used to rank candidate markers. Foldscores ranged from 0 (low) to 1 (high). Foldscore selection threshold demarcated at x ≥ 0.4. DCD had the greatest Foldscore (0.63), followed by alpha-1-acid glycoprotein II (0.58 in Late sequence), Carbonic anhydrase I (0.50), Alpha-1-acid glycoprotein II (0.45 in Early sequence, and S100A8 (0.41).

Carbonic anhydrase I and alpha-1-acid glycoprotein II were excluded after review of the data and literature; the former being a significant component of the erythrocyte proteome that correlated in abundance to the haemoglobin subunit alpha (R^2 0.98, $p < 0.001$) indicative of haemolysis and the latter sharing high homology with the high-abundance alpha-1-acid glycoprotein I that was targeted for immunodepletion. *DCD*, meanwhile, is a multi-product gene that encodes the oncogenic (Y-P30) and pro-cachexic (PIF-core) peptides amongst an anti-microbial peptide (DCD-1) that is secreted in sweat. The peptide sequence quantitated here belongs to the Y-P30 peptide and pro-peptide (Figure 6–2) and is a sequence not shared with other proteins.

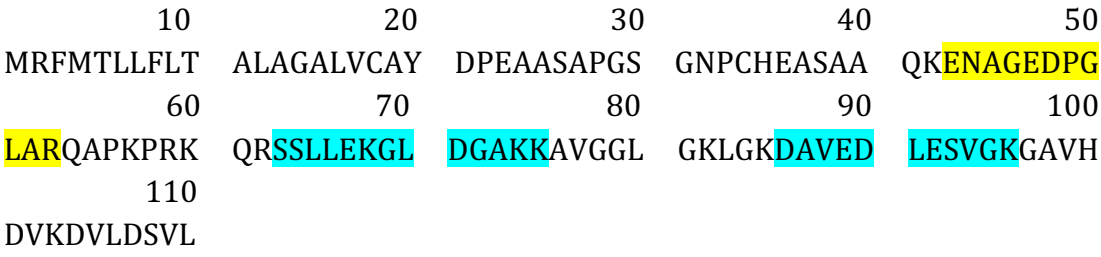


Figure 6–2: *DCD* encodes a 110-AA dermcidin polypeptide consisting of the Y-P30 (AAs 20–49), proteolysis-inducing factor (AAs 20–40), and DCD-1 anti-microbial peptides (AA 63–110). Qualified and quantitated sequences highlighted in blue and yellow, respectively.

Quantitation was based on the relative abundance of six TMT reporter ions revealed in MS/MS spectra after HCD and was compared between time and clinical groups. One and seven HCD spectra were available for DCD and S100A8, respectively (Table 6–3). These showed DCD and S100A8 to have different expression profiles preceding the diagnosis of Dukes A–B cancers; the former being elevated only during the year preceding diagnosis, while S100A8 was raised at 0–1 and 3–4 years versus controls and did not change with tumour growth. DCD levels declined over time in the benign group and controls, while S100A8 remained stable. DCD and S100A8 were identified with low confidence (< 2 unique peptides, < 4 peptide spectral matches) in the Late sequence, and accordingly their behaviour preceding the diagnosis of Dukes C-D cancers is not presented.

Table 6–3: Relative abundance of candidate marker proteins.

Protein	Foldscore	Differential (0–1 year)		Differential (3–4 years)		Longitudinal (0–1 vs 3–4 years)			Spectral counts
		Cancer vs. Control	Cancer vs Benign	Cancer vs. Control	Cancer vs Benign	Control	Benign	Cancer	
Ranked in order of decreasing Foldscore									
DCD	0.63	7.51_1	4.05_1	0.77_1	1.08_1	0.31	0.74	2.89_1	1
Alpha-1-acid glycoprotein II	0.58	1.94_2	2.36_2	0.60_2	0.63_2	0.74	0.68	2.51_2	14
Carbonic anhydrase I	0.50	1.73_2	1.83_2	0.76_2	0.80_2	0.78	0.98	2.14_2	8
Alpha-1-acid glycoprotein II	0.45	2.46_1	2.55_1	1.51_1	1.20_1	1.46	1.04	2.16_1	45
S100A8	0.41	1.83_1	1.29_1	1.77_1	1.74_1	0.86	1.10	0.84_1	7
Ranked in order of decreasing differential expression (cancer vs control) at 0–1 year									
DCD	0.63	7.51_1	4.05_1	0.77_1	1.08_1	0.31	0.74	2.89_1	1
Alpha-1-acid glycoprotein II	0.45	2.46_1	2.55_1	1.51_1	1.20_1	1.46	1.04	2.16_1	45
Alpha-1-acid glycoprotein II	0.58	1.94_2	2.36_2	0.60_2	0.63_2	0.74	0.68	2.51_2	14
S100A8	0.41	1.83_1	1.29_1	1.77_1	1.74_1	0.86	1.10	0.84_1	7
S100A9	0.38	1.76_1	1.12_1	1.65_1	1.79_1	0.72	1.28	0.82_1	11

$_1$ Dukes A–B. $_2$ Dukes C–D. Selection on differential expression between cases and controls at 0–1 year only as an alternative to the Foldscore would have selected S100A9 at the expense of carbonic anhydrase I.

Marker verification

Immunoassays. The complete set of 386 individual sera obtained 0–4 years before diagnosis of benign disease (n 78), cancer (n 154), and from matched controls (n 154) who remained cancer-free during the study period were analysed for DCD and S100A8 levels using commercial ELISA kits. Data were missing for DCD (n 4) owing to insufficient volume. Determinations were made against four-parametric logistic curves run over five 96-well plates (Figure 6–3). Determinations for 195/382 DCD measurements were lower than the lowest standard (0.78 ng mL^{-1}) after correction for dilution but were distinguishable from a background reaction. Constraints on serum volume precluded re-analysis of samples at a lower dilution. For S100A8, three measurements were lower than the lowest calibration standard (50 pg mL^{-1}), while two were higher than the highest calibration standard (12,500 pg mL^{-1}).

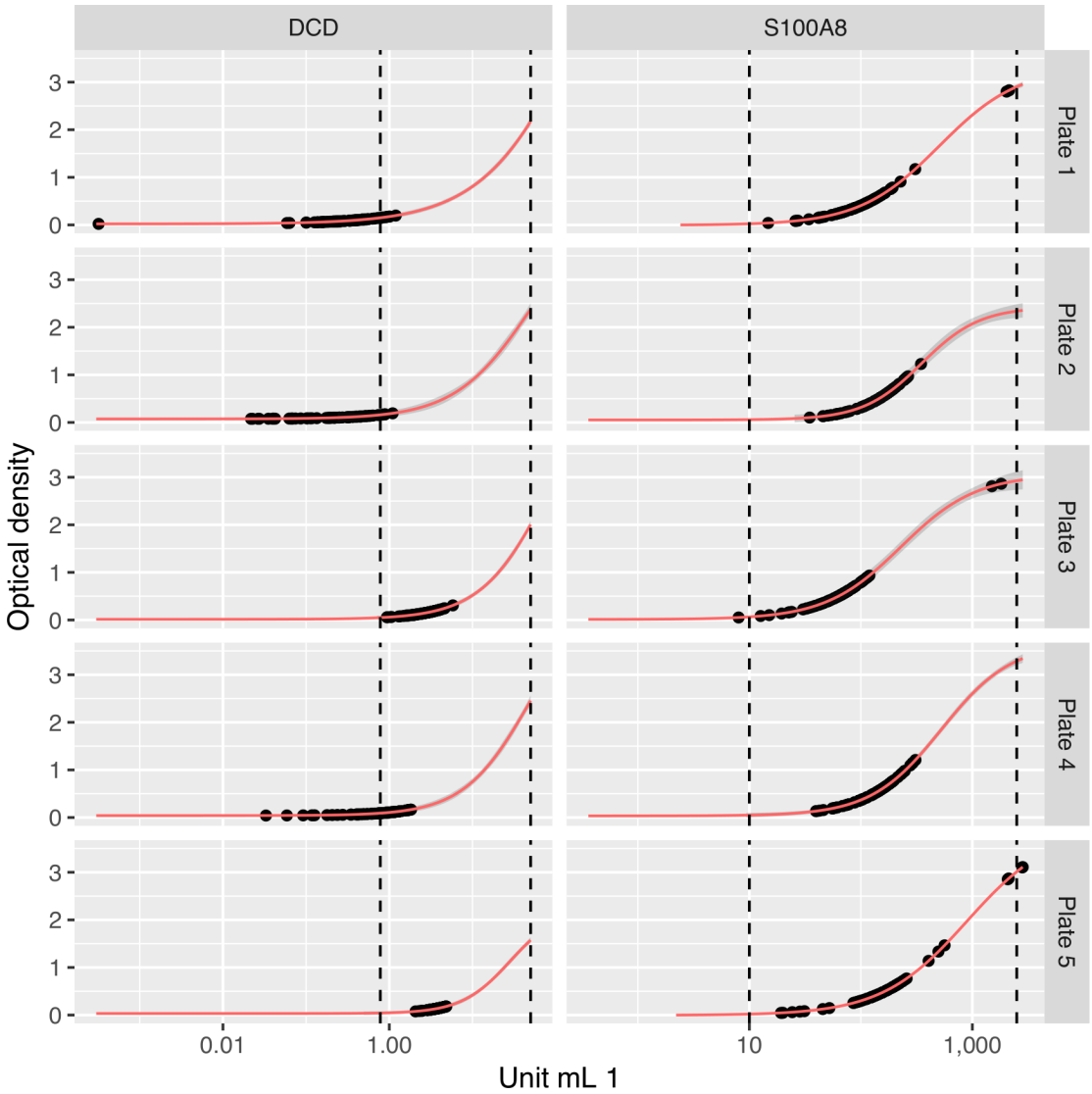


Figure 6–3: Four-parametric logistic curves (red), interpolated samples (black), and 95% confidence intervals (grey). Dashed lines demarcate the upper and lower calibration standards. DCD and S100A8 are expressed as ng mL^{-1} and pg mL^{-1}, respectively.

Differentiation. An overview of serum DCD and S100A8 as a function of time before the diagnoses of benign adenoma or CRC and corresponding levels in women who remained cancer-free is presented in (Figure 6–4). The relative differences in medians (cancer vs. control) at 0–1 and 3-4 years were 0.49 (5.3/10.9 ng mL^{-1}) and 1.1 (7.9/7.2 ng mL^{-1}) for DCD and 0.92 (454.7/492.0 pg mL^{-1}) and 1.06 (625.8/589.8 pg mL^{-1}) for S100A8. DCD was downregulated at 0–1 year before benign adenoma relative to controls (p 0.027). No other comparisons were significant at $p < 0.05$.

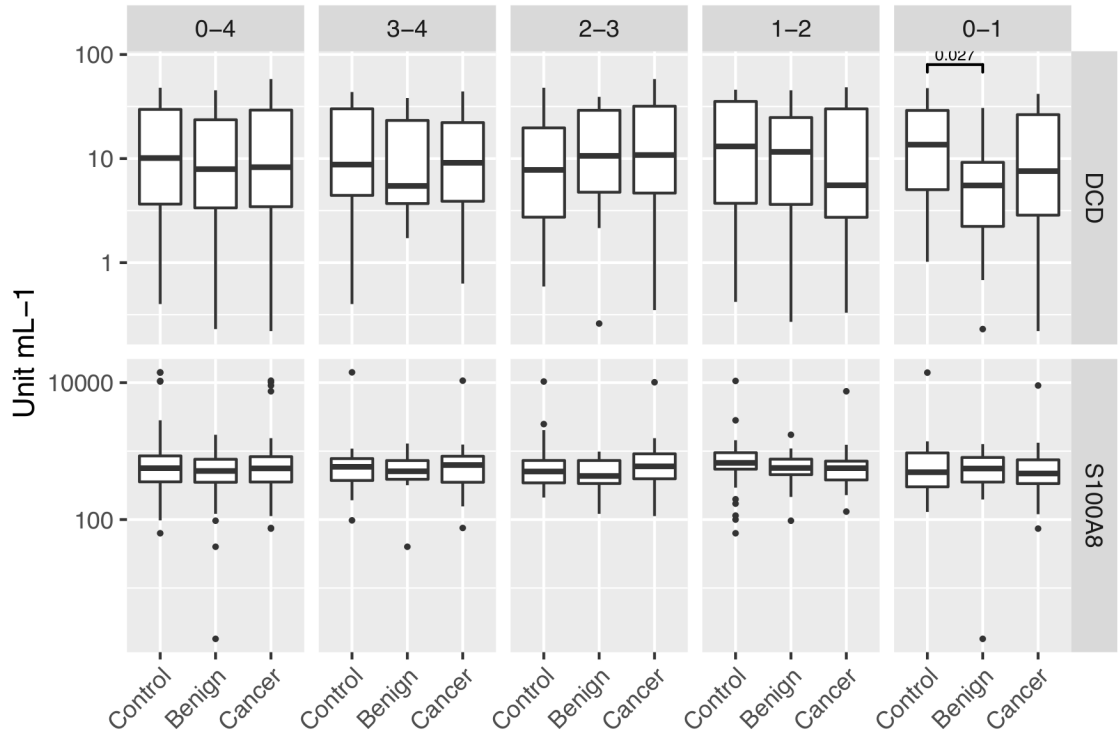


Figure 6–4: DCD and S100A8 levels as a function of time before the diagnosis of benign adenoma or CRC and corresponding levels in matched controls who remained free of cancer. DCD was downregulated in benigns relative to controls at 0–1 year (p 0.027). Boxplots denote the 25/50/75th percentiles, range, and outliers. Time to diagnosis for controls corresponds with that of their matched case. DCD and S100A8 expressed as ng mL^{-1} and pg mL^{-1}, respectively.

Classification. In the absence of significant differences between the groups, the direction of dysregulation for discriminative purposes was defined as that giving the superior AUC for discriminating cancers from controls. For DCD, a threshold of < 0.35 ng mL^{-1} had the best discrimination, at which the specificity and sensitivity were 93% and 9%, respectively. The desired threshold of 95% specificity could not be achieved since ten measurements (> 5% false-positive rate) were 0 ng mL^{-1}. For S100A8, the threshold equivalent to 95% specificity (< 164.5 pg mL^{-1}) gave 6% sensitivity. These thresholds were then applied to samples as a function of lead time afforded (Figure 6–5). At leads times of 0–1, 1–2, 2–3, and 3–4, DCD had sensitivities of 10, 5, 11, and 10%, while S100A8 had 10, 3, 6, 5%.). Notably, DCD was 22–20% sensitive for detecting preclinical benign disease up to 2 years in advance of diagnosis.

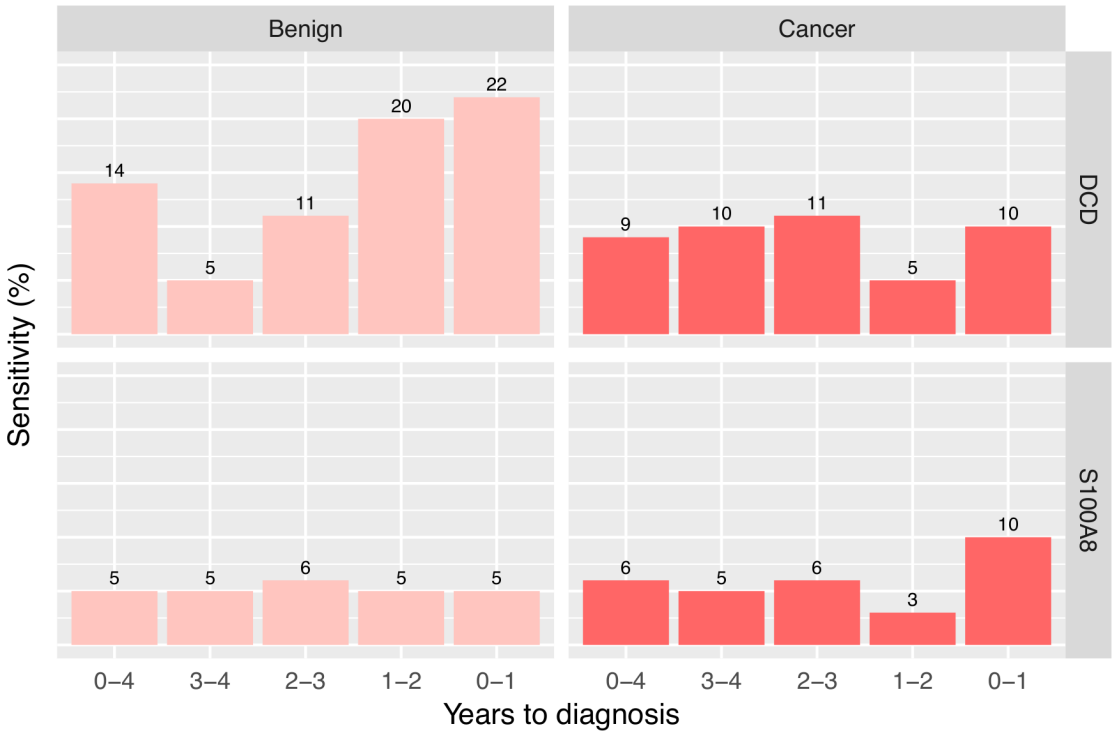


Figure 6–5: Sensitivity of DCD and S100A8 for detecting preclinical benign adenoma and colorectal cancer as a function of time before clinical diagnosis at a threshold equivalent to 95% specificity.

Longitudinal profiles. Spaghetti plots of temporal marker levels in individual women are presented in Figure 6–6. This Figure also highlights that the trend models fitted to all data do not breach the predetermined cut-off thresholds.

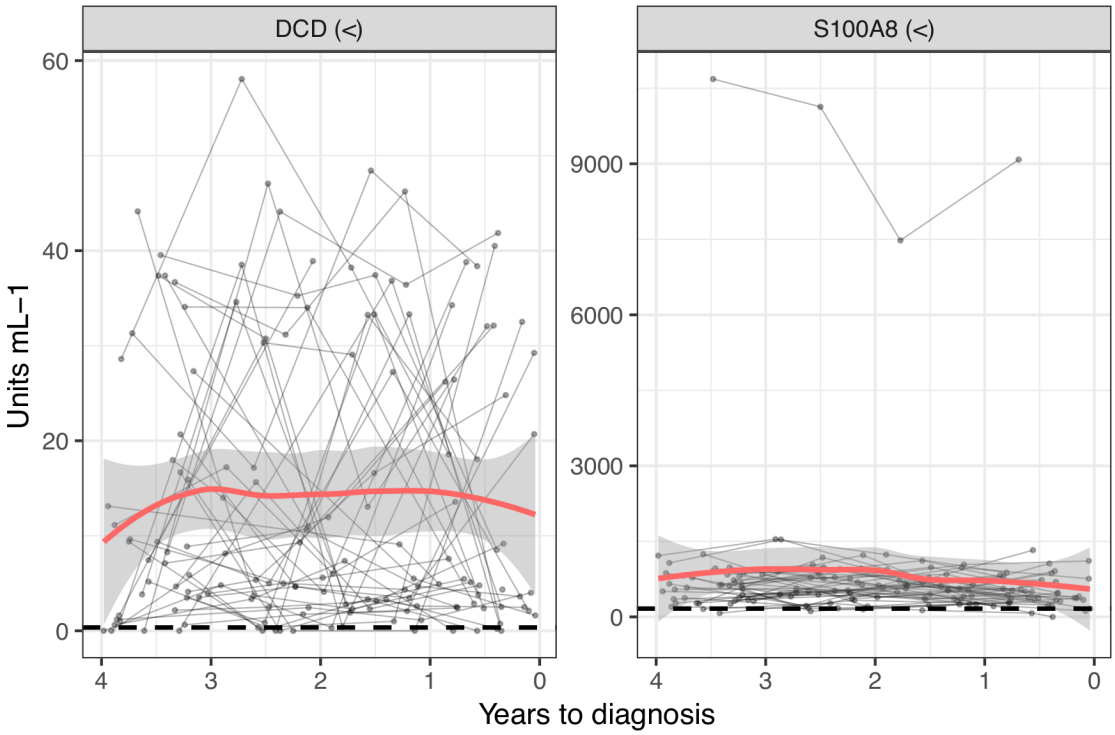


Figure 6–6: Individual marker levels in the lead-up to Dukes A–D cancers and trend profiles based on loess regression (red). The direction of dysregulation (< or >) noted in facet titles. Cut-off thresholds equal to 95% specificity (93% for DCD) demarcated with the dashed line.

Individual variation in marker levels over time was calculated for each woman and are presented as CVs in Table 6–4. Nonetheless, the multivariable analysis dictates that the individual variation did not differ between clinical groups for DCD (p 0.816) or S100A8 (p 0.092).

Table 6–4: Median and range in the variation of individual marker longitudinal profiles.

Marker	Control (n 40)	Benign (n 20)	Dukes A–B (n 20)	Dukes C-D (n 20)
DCD	85 (27–189)	102 (36–200)	92 (39–167)	98 (32–157)
S100A8	41 (17–158)	41 (21–91)	45 (22–84)	33 (15–78)

Variation described as CVs (%).

6.5 Discussion

Principle findings. The present study aimed to identify markers that complement or supersede existing markers previously evaluated. It has been shown herein that serum DCD and S100A8 both detect 10% of preclinical cancers up to a year in advance of diagnosis—and are therefore inferior to CEA and individually would be unsuitable for

screening an asymptomatic population. Monitoring for deviations from individualised baseline profiles for DCD and S100A8 is also unlikely to aide CRC screening.

Strengths and limitations. Many strengths and weaknesses inherent to the study set that were outlined in the previous chapter are also applicable here. Briefly, the strengths were the validity afforded by studying PRoBE-compliant sera, the processing of samples under a standard protocol, and subject matching by age, trial centre, and date of blood donation to limit confounding. Weaknesses were that only sera drawn from postmenopausal women were studied, which has added onus here given that proteomes are gender [23] and hormonal-status dependent [24], and that benigns were not exclusively those most likely to progress to malignancy. Unique to this study were the pros and cons of proteomic techniques. Sample pooling, for example, increases throughout, reduces technical variation but loses all data on individual variation [25]. The removal of high-abundance proteins too, while necessary to detect those lowly abundant (E.g. S100A8 present at ≤ 14 ng mL^{-1}—some seven orders of magnitude lower the omnipresent albumin present at 35–50 mg mL^{-1}), also removes with them their chaperoned proteins that could confound quantitation [26]. DCD and members of the S100 protein family, for example, are detectable in the 'albuminome' [27]. Finally, the lower limit of quantitation for the DCD assay was only marginally higher (0.78 ng mL_{-1}) than the cut-off threshold (< 0.35 ng mL^{-1}) and thus there exists a grey area on the positivity for a marked number of measurements interpolated from absorbencies below the lowest standard. Insufficient sample volume precluded their re-analysis at a lower dilution factor.

Findings in relation to existing literature. Four other studies have, to my knowledge, attempted to uncover candidate markers using preclinical sera—three of these using mass spectrometry. Of these, two used the same Orbitrap XL iteration used here, wherein their analysis of 48 and 364 fractions quantitated 104 [2] and 1,779 proteins [1], respectively. Here, analysis of 200 sample fractions identified 325 proteins, including 135 high-confidence proteins with complete quantitation. The study by Shao used MALDI-TOF/TOF but the overall proteins quantitated were unreported [3]. The quantitative techniques used also differed from label-free quantitation [2] to the chemical labelling of cysteines of intact proteins with acrylamide isotopes [1] or the isobaric tagging of lysines and exposed N-termini of digested peptides used here. Labelling of intact proteins is benefited from being able to resolve the physiochemically diverse proteins (over the more

homogenous peptides) but with the caveat that internal residues are likely inaccessible to these labels [28]. Likewise, labelling of the rarer cysteine residues could reduce sample complexity but is at a trade-off to the quantitative coverage lysines provide. It is thus likely that an amalgamation of various workflows incorporating lysine and cysteine labelling at the protein and peptides levels with varied fractionation techniques (i.e. gel-based) would be beneficial to quantitative coverage. Nonetheless, these studies are far from a comprehensive quantitative coverage of the estimated 10,546 plasma proteins [29] and also highlight the role microarrays could play in marker discovery [4].

All were also not unanimous in the candidates shortlisted; among them being IGFBP2, LRG1, MAPRE1 [1], APOC2, CLU, CO4-B, CO9, FETUA, MASP2, MBL2, PGRP2 [2], ATP1A4, EP300, KAT6A [3], BAG4, CD44, EGFR , IL6ST, and VWF [4]. Those whose preclinical performance were reported were CLU (15% sensitivity at 95% specificity in men only (interpolated from ROC curve)), IL6ST (18% sensitivity/90% specificity), CD44 (19% sensitivity/90% specificity), VWF (25% sensitivity/90% specificity), EGFR (26% sensitivity/90% specificity), and BAG4 (32% sensitivity/90% specificity). (The performances of combinations of these proteins will be reviewed in Chapter 7). These markers were shortlisted as proteins differentially expressed in preclinical cross-sectional samples taken from cases and controls; either on surpassing a threshold set *a priori* (a ≥ or ≤ 1-fold change) [2] or selection via statistical inference tests [1,4] and machine learning [3]. Comparatively we were in a unique position on having measurements for longitudinal samples drawn from women who either developed CRC or remained free of cancer. Thus, the comparisons we were interested in reflected the differential expression between cancers and controls at two time points before diagnoses (0–1 and 3–4 years before diagnosis) and the longitudinal variance in cases and controls. The devised Foldscore was thus an aggregate measure of all four comparisons; designed to reward proteins showing longitudinal variation reflecting the exponential growth of CRCs over time and differential expression between cases and controls, while penalising proteins showing longitudinal variance among controls. This calculation by-and-large shortlisted the same proteins had we considered only differential expression of matched samples taken nearest to diagnosis. DCD and S100A8 were eventually taken forward for validation on their Foldscore and supporting evidence in the literature.

The *DCD* sequence quantitated encodes a Y-P30 peptide that is a demonstrated oncogene [30–32]. It is a major component of the colon tissue proteome and is upregulated two-fold during carcinogenesis [33], but its sensitivity for screening for

cancers < 1 year before diagnosis was 10%. S100A8, meanwhile, is elevated alone or with S100A9 as a calprotectin heterodimer during numerous inflammations, including inflammatory bowel diseases [34,35]. It has a diagnostic sensitivity of 41% [36] but only detects 10% of CRCs a year before diagnosis. Thus, while our data and that reported in the literature suggests DCD and S100A8 to be promising candidates, both fell short of expectations and would be unviable for standalone risk assessment.

Implications for clinicians and policymakers. There is currently insufficient evidence to suggest that evaluating either DCD or S100A8 in prospective screening studies (i.e. beyond phase III) would be worthwhile [37]. Indeed, this study is symptomatic of the discrepancy between the vast literature and the few markers clinically approved [38,39].

Unanswered questions and future research. The vast dynamic range of serum/plasma proteins presents numerous problems for detecting proteins at the trace levels markers leaked or secreted from an asymptomatic tumour are likely to be. Thus, using capture antibodies, protein content was reduced by 99.9% across all ten pools—a magnitude in excess of the 97% attributed to these proteins, however. Without further experiments, it is only possible to speculate that those unrecovered proteins were those smaller than the 5 kDa-pore ultrafilter membranes (thymosin beta (5kDa) was the smallest protein identified), or those bound to the polyether consumables, capture antibodies or targeted proteins. Since targeted proteins still featured in the protein lists obtained, non-specific removal is most likely. Nonetheless, targeted features were among the most scanned analytes, highlighting the ever-present problem these proteins pose to marker discovery. The use of Michaelis-Menten kinetics to digest target proteins upstream immunodepletion may be worthwhile [40].

No individual markers of the eight verified in the present or preceding chapters have the required performance to replace the gFOBT or FIT. It was posited in Chapter 5 that the barrier to regulatory-approved markers for cancer screening is for the lack of preclinical samples available for discovery work. This belief was founded on the discordance between the prolificacy in which novel markers are reported and their failure to reach the clinic [38]. On closer study, the barriers to clinical translation are challenges of a technical or biological nature. Modelling the math of marker secretion and/or shedding into circulation and its clearance and excretion, it was concluded that detection of an asymptomatic solid cancer of $1mm^3$ would be possible only with a wholly tumour-

specific marker with a secretion/shedding rate substantially higher than empirically observed hitherto or via significant improvements to the analytical sensitivity of marker assays [41]. In the meanwhile, future work should establish the performance of a complementary panel of two or more of these proteins for detecting asymptomatic CRC [42].

6.6 Conclusions

DCD and S100A8 lack sufficient sensitivity for screening for CRC in asymptomatic, average-risk women.

6.7 References

[1] J.J. Ladd, T. Busald, M.M. Johnson, Q. Zhang, S.J. Pitteri, H. Wang, D.E. Brenner, P.D. Lampe, R. Kucherlapati, Z. Feng, R.L. Prentice, S.M. Hanash, Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women, Cancer Prev. Res. (Phila). 5 (2012) 655–64. doi:10.1158/1940-6207.CAPR-11-0412.

[2] M. Bertuzzi, C. Marelli, R. Bagnati, A. Colombi, R. Fanelli, C. Saieva, M. Ceroti, B. Bendinelli, S. Caini, L. Airoldi, D. Palli, Plasma clusterin as a candidate pre-diagnosis marker of colorectal cancer risk in the Florence cohort of the European Prospective Investigation into Cancer and Nutrition: a pilot study, BMC Cancer. 15 (2015) 56. doi:10.1186/s12885-015-1058-7.

[3] S. Shao, B.A. Neely, T.-C. Kao, J. Eckhaus, J. Bourgeois, J. Brooks, E.E. Jones, R.R. Drake, K. Zhu, Proteomic Profiling of Serial Prediagnostic Serum Samples for Early Detection of Colon Cancer in the U.S. Military, Cancer Epidemiol. Biomarkers Prev. 26 (2017) 711–718. doi:10.1158/1055-9965.EPI-16-0732.

[4] J. Rho, J.J. Ladd, C.I. Li, J.D. Potter, Y. Zhang, D. Shelley, D. Shibata, D. Coppola, H. Yamada, H. Toyoda, T. Tada, T. Kumada, D.E. Brenner, S.M. Hanash, P.D. Lampe, Protein and glycomic plasma markers for early detection of adenoma and colon cancer, Gut. 67 (2018) 473–484. doi:10.1136/gutjnl-2016-312794.

[5] N.L. Anderson, N.G. Anderson, The human plasma proteome: history, character, and diagnostic prospects, Mol. Cell. Proteomics. 1 (2002) 845–67.

[6] G.L. Hortin, D. Sviridov, The dynamic range problem in the analysis of the plasma proteome, J. Proteomics. 73 (2010) 629–36. doi:10.1016/j.jprot.2009.07.001.

[7] T. Liu, W.-J. Qian, H.M. Mottaz, M.A. Gritsenko, A.D. Norbeck, R.J. Moore, S.O. Purvine, D.G. Camp, R.D. Smith, Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry, Mol. Cell. Proteomics. 5 (2006) 2167–74. doi:10.1074/mcp.T600039-MCP200.

[8] B. Manadas, V.M. Mendes, J. English, M.J. Dunn, Peptide fractionation in proteomics approaches, Expert Rev. Proteomics. 7 (2010) 655–663. doi:10.1586/epr.10.46.

[9] H. Steen, M. Mann, The ABC's (and XYZ's) of peptide sequencing, Nat. Rev. Mol. Cell Biol. 5 (2004) 699–711. doi:10.1038/nrm1468.

[10] R. Aebersold, M. Mann, Mass spectrometry-based proteomics, Nature. 422 (2003) 198–207. doi:10.1038/nature01511.

[11] P. Roepstorff, J. Fohlman, Proposal for a common nomenclature for sequence ions in mass spectra of peptides, Biol. Mass Spectrom. 11 (1984) 601–601. doi:10.1002/bms.1200111109.

[12] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis. 20 (1999) 3551–67. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

[13] A. Thompson, J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A.K.A. Mohammed, C. Hamon, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, Anal. Chem. 75 (2003) 1895–904. doi:10.1021/ac0262560.

[14] G.C. McAlister, E.L. Huttlin, W. Haas, L. Ting, M.P. Jedrychowski, J.C. Rogers, K. Kuhn, I. Pike, R.A. Grothe, J.D. Blethrow, S.P. Gygi, Increasing the Multiplexing Capacity of TMTs Using Reporter Ion Isotopologues with Isobaric Masses, Anal. Chem. 84 (2012) 7469–7478. doi:10.1021/ac301572t.

[15] U. Menon, A. Gentry-Maharaj, A. Ryan, A. Sharma, M. Burnell, R. Hallett, S. Lewis, A. Lopez, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, S. Skates, M. Parmar, I. Jacobs, Recruitment to multicentre trials—lessons from UKCTOCS: descriptive study, BMJ. 337 (2008) a2079–a2079. doi:10.1136/bmj.a2079.

[16] U. Menon, A. Gentry-Maharaj, R. Hallett, A. Ryan, M. Burnell, A. Sharma, S. Lewis, S. Davies, S. Philpott, A. Lopes, K. Godfrey, D. Oram, J. Herod, K. Williamson, M.W. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N.N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, N. Singh, A. Dawnay, S.J. Skates, M. Parmar, I. Jacobs, Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), Lancet. Oncol. 10 (2009) 327–40. doi:10.1016/S1470-2045(09)70026-9.

[17] R Core Team, R: a language and environment for statistical computing, (2016). https://www.r-project.org/.

[18] C. Ritz, F. Baty, J.C. Streibig, D. Gerhard, Dose-Response Analysis Using R, PLoS One. 10 (2015) e0146021. doi:10.1371/journal.pone.0146021.

[19] M. Stevenson, T. Nunes, C. Heuer, J. Marshall, J. Sanchez, R. Thornton, J. Reiczigel, J. Robison-Cox, P. Sebastiani, P. Solymos, K. Yoshida, G. Jones, S. Pirikahu, S. Firestone, R. Kyle, J. Popp, M. Jay, epiR: tools for the Analysis of Epidemiological Data, (2017). https://cran.r-project.org/package=epiR.

[20] H. Wickham, ggplot2: elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.

[21] M. López-Ratón, M.X. Rodríguez-Álvarez, C.C. Suárez, F.G. Sampedro, OptimalCutpoints: an R Package for Selecting Optimal Cutpoints in Diagnostic Tests, J. Stat. Softw. 61 (2014) 1–36. doi:10.18637/jss.v061.i08.

[22] H. Wickham, L. Henry, tidyr: easily Tidy Data with "spread()" and "gather()" Functions, (2018). https://cran.r-project.org/package=tidyr.

[23] K. Miike, M. Aoki, R. Yamashita, Y. Takegawa, H. Saya, T. Miike, K. Yamamura, Proteome profiling reveals gender differences in the composition of human serum, Proteomics. 10 (2010) 2678–2691. doi:10.1002/pmic.200900496.

[24] H. Katayama, S. Paczesny, R. Prentice, A. Aragaki, V.M. Faca, S.J. Pitteri, Q. Zhang, H. Wang, M. Silva, J. Kennedy, J. Rossouw, R. Jackson, J. Hsia, R. Chlebowski, J. Manson, S. Hanash, Application of serum proteomics to the Women's Health Initiative conjugated equine estrogens trial reveals a multitude of effects relevant to clinical findings, Genome Med. 1 (2009) 47. doi:10.1186/gm47.

[25] A.P. Diz, M. Truebano, D.O.F. Skibinski, The consequences of sample pooling in proteomics: an empirical study, Electrophoresis. 30 (2009) 2967–2975. doi:10.1002/elps.200900210.

[26] J. Granger, J. Siddiqui, S. Copeland, D. Remick, Albumin depletion of human plasma also removes low abundance proteins including the cytokines, Proteomics. 5 (2005) 4713–8. doi:10.1002/pmic.200401331.

[27] Z. Liu, S. Li, H. Wang, M. Tang, M. Zhou, J. Yu, S. Bai, P. Li, J. Zhou, P. Xie, Proteomic and network analysis of human serum albuminome by integrated use of quick crosslinking and two-step precipitation, Sci. Rep. 7 (2017) 9856. doi:10.1038/s41598-017-09563-w.

[28] J. Sinclair, J.F. Timms, Quantitative profiling of serum samples using TMT protein labelling, fractionation and LC-MS/MS, Methods. 54 (2011) 361–9. doi:10.1016/j.ymeth.2011.03.004.

[29] V. Nanjappa, J.K. Thomas, A. Marimuthu, B. Muthusamy, A. Radhakrishnan, R. Sharma, A. Ahmad Khan, L. Balakrishnan, N.A. Sahasrabuddhe, S. Kumar, B.N. Jhaveri, K.V. Sheth, R. Kumar Khatana, P.G. Shaw, S.M. Srikanth, P.P. Mathur, S. Shankar, D. Nagaraja, R. Christopher, S. Mathivanan, R. Raju, R. Sirdeshmukh, A. Chatterjee, R.J. Simpson, H.C. Harsha, A. Pandey, T.S.K. Prasad, Plasma Proteome Database as a resource for proteomics research: 2014 update, Nucleic Acids Res. 42 (2014) D959-65. doi:10.1093/nar/gkt1251.

[30] D. Porter, S. Weremowicz, K. Chin, P. Seth, A. Keshaviah, J. Lahti-Domenici, Y.K. Bae, C.L. Monitto, A. Merlos-Suarez, J. Chan, C.M. Hulette, A. Richardson, C.C. Morton, J. Marks, M. Duyao, R. Hruban, E. Gabrielson, R. Gelman, K. Polyak, A neural survival factor is a candidate oncogene in breast cancer, Proc. Natl. Acad. Sci. 100 (2003) 10931–10936. doi:10.1073/pnas.1932980100.

[31] A.G. Lowrie, S.J. Wigmore, D.J. Wright, I.D. Waddell, J.A. Ross, Dermcidin expression in hepatic cells improves survival without N-glycosylation, but requires asparagine residues, Br. J. Cancer. 94 (2006) 1663–71. doi:10.1038/sj.bjc.6603148.

[32] G.D. Stewart, A.G. Lowrie, A.C.P. Riddick, K.C.H. Fearon, F.K. Habib, J.A. Ross, Dermcidin expression confers a survival advantage in prostate cancer cells subjected to oxidative stress or hypoxia, Prostate. 67 (2007) 1308–1317. doi:10.1002/pros.20618.

[33] J.R. Wiśniewski, P. Ostasiewicz, K. Duś, D.F. Zielińska, F. Gnad, M. Mann, Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma, Mol. Syst. Biol. 8 (2012) 611. doi:10.1038/msb.2012.44.

[34] D. Foell, M. Frosch, C. Sorg, J. Roth, Phagocyte-specific calcium-binding S100 proteins as clinical laboratory markers of inflammation, Clin. Chim. Acta. 344 (2004) 37–51. doi:10.1016/j.cccn.2004.02.023.

[35] M.R. Konikoff, L.A. Denson, Role of fecal calprotectin as a biomarker of intestinal inflammation in inflammatory bowel disease, Inflamm. Bowel Dis. 12 (2006) 524–34.

[36] H.J. Kim, H.J. Kang, H. Lee, S.T. Lee, M.H. Yu, H. Kim, C. Lee, Identification of S100A8 and S100A9 as serological markers for colorectal cancer, J Proteome Res. 8 (2009) 1368–1379. doi:10.1021/pr8007573.

[37] M.S. Pepe, R. Etzioni, Z. Feng, J.D. Potter, M.L. Thompson, M. Thornquist, M. Winget, Y. Yasui, Phases of biomarker development for early detection of cancer, J. Natl. Cancer Inst. 93 (2001) 1054–61.

[38] E.P. Diamandis, The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem?, BMC Med. 10 (2012) 87. doi:10.1186/1741-7015-10-87.

[39] S.E. Kern, Why Your New Cancer Biomarker May Never Work: Recurrent Patterns and Remarkable Diversity in Biomarker Failures, Cancer Res. 72 (2012) 6097–6101. doi:10.1158/0008-5472.CAN-12-3232.

[40] B.R. Fonslow, B.D. Stein, K.J. Webb, T. Xu, J. Choi, S.K. Park, J.R. Yates, Digestion and depletion of abundant proteins improves proteomic coverage, Nat. Methods. 10 (2013) 54–6. doi:10.1038/nmeth.2250.

[41] S.S. Hori, S.S. Gambhir, Mathematical model identifies blood biomarker-based early cancer detection strategies and limitations, Sci. Transl. Med. 3 (2011) 109ra116. doi:10.1126/scitranslmed.3003110.

[42] C.A.K. Borrebaeck, Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer, Nat. Rev. Cancer. 17 (2017) 199–204. doi:10.1038/nrc.2016.153.

Chapter 7. Exploring the Use of Multi-Marker Panels to Enhance the Performance of Individual Markers

A Retrospective Case-Control Study Nested Within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS).

7.1 Background

The preceding chapters demonstrated that neither CA125, CEA, CYFRA21-1, DCD, FAP, S100A8, TIMP-1, nor VEGFA has the required sensitivity to be of value for standalone screening for CRC in asymptomatic, average-risk women. On rationale that CRCs vary in their clinical, histological, and molecular subtyping, it is posited that a panel of complementary markers may improve upon the sensitivity of an individual marker. Concluding this thesis, the present chapter aims to determine if a combination of the current markers has the required performance for detecting preclinical CRC in asymptomatic, average-risk women.

7.2 Methods

The subjects and sample measurements studied herein are those studied in Chapters 5 (CA125, CEA, CYFRA21-1, TIMP-1, VEGFA) and 6 (DCD, S100A8). These were case-controls studies nested within the multimodal arm of the UKCTOCS involving 50,640 women who donated blood annually. Full methods are described previously and are briefly outlined below.

Subjects. One-hundred subjects were studied, including 40 cases, 20 matched benigns, and 40 matched controls. Cases were clinically verified adenocarcinomas of the colon or rectum, excluding those of the appendix and those with a signet-ring cell or mucinous morphology. Benigns were women with a diagnosis of benign adenoma (D12, excluding those of the appendix (D12.1), anus and anal canal (D12.9)) recorded in HES and no cancer diagnoses of any origin according to cancer and death registries, HES, or self-reporting. Controls were women who had no diagnosis of cancer of any origin according to cancer and death registries, HES, and self-reporting as of 1 September 2013. The baseline characteristics of subjects and clinical and histological characteristics of benign adenomas and adenocarcinomas were presented previously (Table 5–5 & Table 5–6).

Samples. There were 386 serum samples (154 controls, 78 benign, 154 cancers (Table 5–2)) drawn from 100 subjects (40 controls, 20 beings, 40 cancers). Each woman had 3–4 samples eligible for the study (mean 3.9). Samples predated the diagnosis of benign adenoma or CRC by 0–4 years.

Immunoassays. Sera CA125, CEA, and CYFRA21-1 were measured using commercialised Cobas immunoassays approved for diagnostic use (Chapter 5). All other markers were measured using commercial immunoassay kits outlined in Chapters 5 (FAP, TIMP-1, VEGFA) and 6 (DCD, S100A8). CA125 is presented as U mL^{-1}, CEA, CYFRA21-1, DCD, FAP, & TIMP-1 as ng mL^{-1}, and S100A8 & VEGFA as pg mL^{-1}.

Data analyses

Multi-marker panels. Multi-marker panels of all 248 combinations of 2–8 markers ((8^2) minus eight individual markers verified previously) were modelled using logistic regression after exclusion of missing data. The outcome variable was a verified CRC diagnosis or no cancer diagnosis during the study period (defined as 'cases' and 'controls', respectively). Predictor variables were sera CA125, CEA, CYFRA21-1, DCD, FAP, S100A8, TIMP-1, and VEGFA levels on a continuous scale.

Ranking. Models were ranked from high to low in increasing order of adjusted Akaike's Information Criterion (AICc), and the performance of the top-three with the lowest AICc was assessed. Briefly, AIC quantitates how well a statistical model fits the data after penalisation for model complexity.

Classification. A composite score of panel members was derived for all samples from the linear predictor form of the logistic regression equation wherein the probability of CRC given a set of marker levels (*P(Y|xi)*) is calculated as:

$$P(Y \mid xi) = \frac{1}{1 + e^{-(\beta 0 + \beta 1X1 + \beta 2X2 + \beta nXn)}} \quad \textbf{(1)}$$

Given the computed intercept (*β0*), regression coefficients (*βn)*, and marker levels (*Xn*), the linear predictor was calculated as:

$$\beta 0 + \beta 1X1 + \beta 2X2 + \beta nXn \quad \textbf{(2)}$$

Cut-off thresholds were defined as the linear predictor value equivalent to 95% specificity (i.e. a 5% false-positive rate). The sensitivity for detecting preclinical benign adenoma and CRC as a function of time before diagnosis was thereafter calculated at this threshold with 95% confidence intervals. The Receiver Operating Characteristic (ROC) was fitted for each selected panel and the Area under the Curve (AUC), and 95% confidence intervals were thereafter calculated. The AUC in this context describes the probability that marker levels are higher in a randomly sampled case than in a random control.

Statistical analyses. All analyses were made with R version 3.3.2 [1] running epiR [2], ggplot2 [3], glmulti [4], OptimalCutpoints [5], pROC [6], and tidyr [7] packages. A p value < 0.05 was considered statistically significant.

7.3 Results

Multi-marker panels. Six samples had incomplete data for all eight markers (1.9% (6/308)). Excluding these was not thought to introduce bias. Thus, panels were trained on eight marker measurements obtained for 149 serial sera drawn from 40 women who developed CRC and 153 sera from 40 matched controls (2,416 measurements from 302 samples in total) and tested on 378 serial sera (training set plus 76 sera from 20 women who developed benign adenoma).

Ranking. The AICc for all 256 marker combinations (including eight panels of individual markers presented previously) ranged from 407.1–428.3 about a median of 418.1 (Figure 7–1). The three highest-ranked panels (termed panels I–III inclusively) were in the 99th percentile with AICcs $\geq$ 427. Panel I consisted of CA125, CEA, FAP, TIMP-1, & VEGFA; Panel II of CA125, CEA, FAP, & VEGFA; and Panel III of CA125, CEA, & VEGFA.

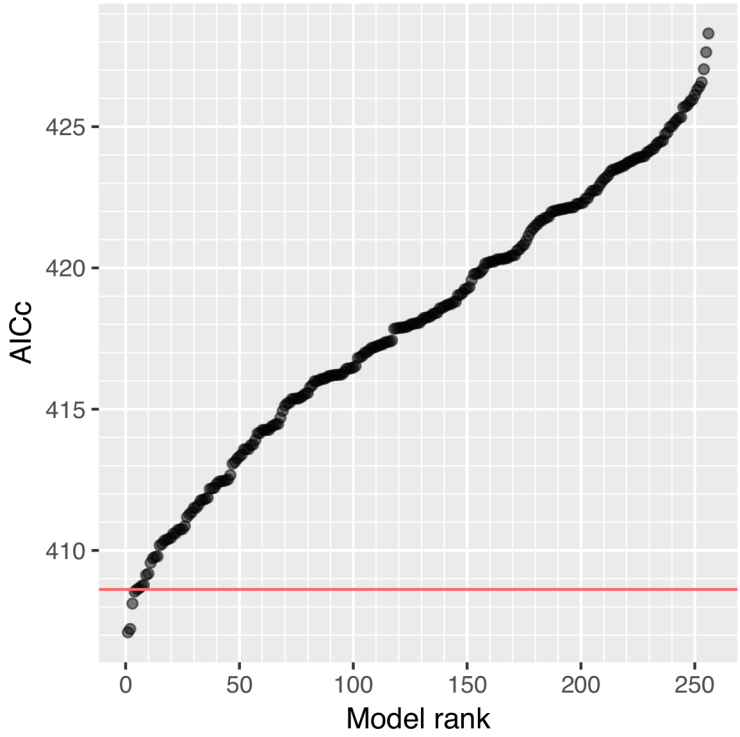


Figure 7–1: Model rank versus adjusted Akaike's Information Criterion. The top-three ranked panels are separated by the red horizontal line.

Classification. The sensitivity for detecting CRC at lead times of 0–4 and 0–1 years were 0.10 (95% CI 0.06–0.16) and 0.23 (95% CI 0.11–0.38) for of Panel I (CA125, CEA, FAP, TIMP-1, & VEGFA), 0.09 (95% CI 0.05–0.14) and 0.17 (95% CI 0.07–0.33) for Panel II (CA125, CEA, FAP, and VEGFA), and 0.07 (95% CI 0.03–0.12) and 0.17 (95% CI 0.07–0.33) for Panel III (CA125, CEA, and VEGFA). The sensitivity for detecting benign adenoma at lead times of 0–4 and 0–1 year(s) were 0.04 (95% CI 0.01–0.11) and 0.00 (95% CI 0.00–0.19) for Panel I, 0.05 (95% CI 0.01–0.13) and 0.06 (95% CI 0.00–0.27) for Panel II, and 0.00 (95% CI 0.00–0.05) and 0.00 (95% CI 0.00–0.19) for Panel III. The sensitivities as a function of lead time are presented in Figure 7–2.

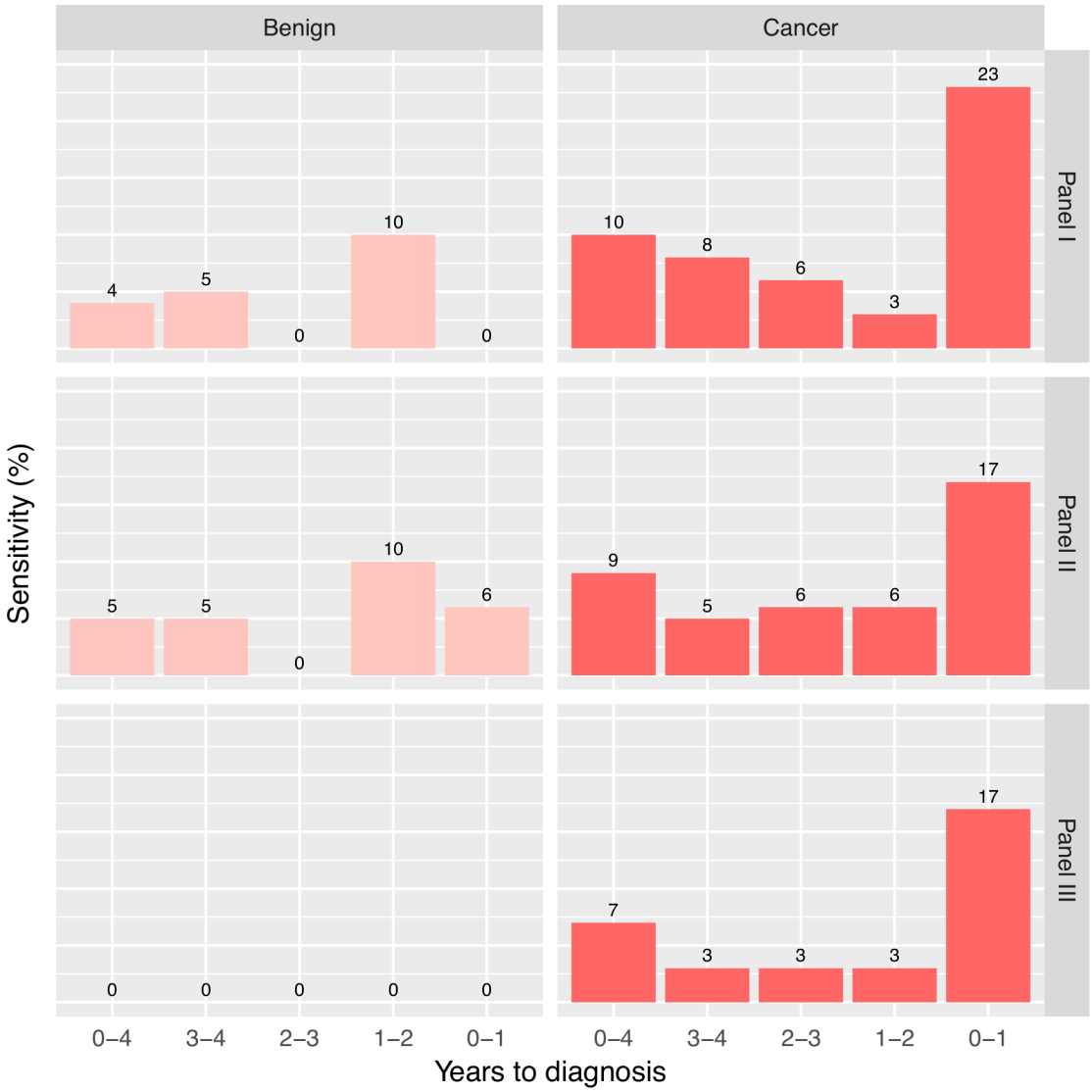


Figure 7–2: Sensitivity of the three highest ranked multi-marker panels for detecting preclinical benign adenoma and colorectal cancer as a function of time before clinical diagnosis at a threshold equivalent to 95% specificity.

Panel I (CA125, CEA, FAP, TIMP-1, VEGFA) had the greatest AUC of 0.66 (95% CI 0.60–0.72), followed by Panel II (CA125, CEA, FAP, VEGFA) and III (CA125, CEA, VEGFA) with AUCs of 0.64 (95% CI 0.58–0.71) and 0.64 (95% CI 0.57–0.70), respectively. The corresponding AUCs at 0–1 year before diagnosis were 0.75 (95% CI 0.66–0.84), 0.73 (95% CIs 0.64–0.82), and 0.70 (95% CIs 0.61–0.79) for Panels I–III inclusively.

7.4 Discussion

Principle findings. A panel of multiple markers is in many circumstances superior to any individual marker. Herein, however, the highest-ranked panel of CA125, CEA, FAP,

TIMP-1, and VEGFA performed no better than CEA alone. The panel had a sensitivity of 23% for detecting preclinical CRCs up to one year before clinical diagnosis and an AUC of 0.75. The sensitivity and AUC including measurements up to 4 years before diagnosis was 10% and 0.66, respectively.

Strengths and limitations. The study is strengthened by the exhaustive manner in which all 248 permutations of 2–8 marker combinations were assessed. Furthermore, the adjusted Akaike's Information Criterion was apposite for ranking panels given its correlation with the AUC where calculated. The suitability of using AICs as a ranking metric could be explored further given sufficient computation and time. There are also several limitations. Logistic regression, for example, assumes that all observations are independent of one another; when in reality there were 3–4 repeated measurements per woman, each reflecting an individualised baseline profile. Time permitted, a multilevel model accounting for the hierarchical nature of longitudinal data should be fitted. Secondly, there is a potential of overfitting the model to the current sample measurements, which would bias performance estimates derived by apparent validation towards overestimation. Without external validation on an independent set of preclinical samples, panels could be assessed on iterations of bootstrap resamples [8], with or without recalibration or shrinkage of overfitted coefficients [9]. Nonetheless, the UKCTOCS biobank stores many cross-sectional should future validation on an independent sample set be necessary.

Findings in relation to existing literature. All four of the PRoBE studies introduced in Chapter 6 modelled multi-marker panels with candidates shortlisted from their analyses of preclinical sera. To add to these is another that follows-on from the study by Wild whose estimates on the diagnostic performances for many proteins were presented in Chapter 5 [10]. Only CEA and FAP were candidates shared with two other panels. In the first, testing on an independent yet internal set of preclinical samples from the same biobank, a panel of CEA, IGFBP2, LRG1, and MAPRE1 had a sensitivity of 41% at 95% specificity (AUC 0.60) for detecting CRCs up to seven months before diagnosis [11]. Our panel, by comparison, had a sensitivity of 27% when restricted to samples with a lead time of seven months or less. The study also reports that the addition of IGFBP2, LRG1, and MAPRE1 enhanced the sensitivity of CEA alone, which had a sensitivity of 19% at the same specificity (AUC 0.60). Addition of our markers did not improve on the

sensitivity of CEA alone. In the second study, a panel of CEA, anti-p53, FAP, ferritin, and OPN trained on diagnostic measurements had a sensitivity of 42% at 95% specificity and an AUC of 0.78 when validated on an independent set of samples taken before screening colonoscopy [12]. A limitation to testing in a screening setting, however, is that the lead time afforded cannot be established. Nonetheless, when stratified to subjects who had both have blood and stool samples obtained before colonoscopy, the study demonstrated the panel to be as sensitive as the gFOBT (39%) but inferior to the FIT (78%).

The three panels reported in the literature include markers that were unstudied here. A panel derived of the mass-spectra peaks identified as ATP1A4, EP300, KAT6A, and an unidentified protein had a sensitivity of 69% at 67% specificity for detecting disease up to 1 year in advance of diagnosis [13], though it is unclear how this panel would perform at a specificity threshold viable within a screening context. It is for this reason that marker sensitivity has been determined at a threshold equivalent to 95% specificity throughout this thesis—not because of this false-positive rate has particular relevance to the prevalence of CRC or endoscopic capacity of the NHS but because it is benchmark widely adopted by the literature. Elsewhere, a panel of BAG4, CD44, IL6ST, and VWF at 90% specificity had a sensitivity of 42% (AUC 0.79) when applied to preclinical samples drawn up to three-years before diagnosis [14]. As are the limitations of our study, the apparent validation on the same measurements used for model derivation likely renders these estimates as optimistic. Finally, reporting on an internal validation using an independent preclinical sample set drawn from the same biobank, CLU alone was superior (AUC 0.72) to a panel of CLU, APOC2, and MASP2 (AUC 0.62); though sensitivity estimates and lead-times were unreported [15]. None of these panels have been validated on an independent sample set.

Implications for clinicians and policymakers. The arduous quest for FDA approval or CE marking of novel screening tests is exemplified by how few have been successful—until the recent pre-market approval of Cologuard and Epi proColon 2.0 for screening average-risk persons aged 50 years and over. Cologuard (Exact Sciences) is a multitarget test that looks for cancer abnormalities exfoliated onto stool samples. The test combines real-time quantitative PCR assays for mutant *KRAS* and hypermethylation *BMP3* & *NDRG4* with a haemoglobin immunoassay. Cologuard has, in comparison to the FIT, greater sensitivity (92–100% vs 74–75%) and AUC (0.94 vs. 0.89) but lower specificity

for screening (87–90 % vs. 95–96%) [16,17]. Epi proColon (Epigenomics), meanwhile, is a qualitative real-time PCR that detects hypermethylated *SEPT9* (m*SEPT9*) in cell-free DNA circulating in plasma. The sensitivity of Epi proColon was 48% (35, 63, 46, and 77% for Dukes A–D inclusively) at 92% specificity [18]. While inferior in specificity to the FIT [19], it may nonetheless have uses for improving the uptake of those who avert conventional screening [20]. By comparison, the five-protein panel assessed herein has inferior performance and does to attain the precedent set by the American and European regulatory bodies. Furthermore, CEA plus CA125, FAP, TIMP-1, and VEGFA is no more sensitive than CEA alone for detecting CRCs within one year of diagnosis—and is worse still for detecting at a lead time greater than one year. The current evidence, therefore, implies that it would not be worthwhile taking this panel forward for evaluation in an independent set of samples.

Unanswered questions and future research. The additional logistics and expenditure being time or costs for assaying an additional four proteins over CEA alone are not justified. Accordingly, the challenge going forward is to enrich the sensitivity of CEA alone. Exact Science's Epi proColon is the first and (as of late 2018) only liquid biopsy approved for CRC screening (performance discussed above). Future work could explore whether the m*SEPT9* assay could enhance the performance of CEA—or vice versa if CEA could supplement the specificity of m*SEPT9*. A preliminary study of high-risk patients suggests this may be a worthwhile exploration [21].

7.5 Conclusions

A multi-marker panel of CA125, CEA, FAP, TIMP-1 and VEGFA is not suitable for screening for CRC. At 95% specificity, it had a sensitivity of 23% for detecting CRCs up to one-year and 10% up to four years before diagnosis. The panel performed no better than CEA alone.

7.6 References

[1] R Core Team, R: a language and environment for statistical computing, (2016). https://www.r-project.org/.

[2] M. Stevenson, T. Nunes, C. Heuer, J. Marshall, J. Sanchez, R. Thornton, J. Reiczigel, J. Robison-Cox, P. Sebastiani, P. Solymos, K. Yoshida, G. Jones, S. Pirikahu, S. Firestone, R. Kyle, J. Popp, M. Jay, epiR: tools for the Analysis of Epidemiological Data, (2017). https://cran.r-project.org/package=epiR.

[3] H. Wickham, ggplot2: elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.

[4] V. Calcagno, glmulti: Model selection and multimodel inference made easy, (2013). https://cran.r-project.org/package=glmulti.

[5] M. López-Ratón, M.X. Rodríguez-Álvarez, C.C. Suárez, F.G. Sampedro, OptimalCutpoints: an R Package for Selecting Optimal Cutpoints in Diagnostic Tests, J. Stat. Softw. 61 (2014) 1–36. doi:10.18637/jss.v061.i08.

[6] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinformatics. 12 (2011) 77. doi:10.1186/1471-2105-12-77.

[7] H. Wickham, L. Henry, tidyr: easily Tidy Data with "spread()" and "gather()" Functions, (2018). https://cran.r-project.org/package=tidyr.

[8] E.W. Steyerberg, F.E. Harrell, G.J. Borsboom, M.J. Eijkemans, Y. Vergouwe, J.D. Habbema, Internal validation of predictive models: efficiency of some procedures for logistic regression analysis, J. Clin. Epidemiol. 54 (2001) 774–81.

[9] E.W. Steyerberg, G.J.J.M. Borsboom, H.C. van Houwelingen, M.J.C. Eijkemans, J.D.F. Habbema, Validation and updating of predictive logistic regression models: a study on sample size and shrinkage, Stat. Med. 23 (2004) 2567–2586. doi:10.1002/sim.1844.

[10] N. Wild, H. Andres, W. Rollinger, F. Krause, P. Dilba, M. Tacke, J. Karl, A combination of serum markers for the early detection of colorectal cancer, Clin Cancer Res. 16 (2010) 6111–6121. doi:10.1158/1078-0432.CCR-10-0119.

[11] J.J. Ladd, T. Busald, M.M. Johnson, Q. Zhang, S.J. Pitteri, H. Wang, D.E. Brenner, P.D. Lampe, R. Kucherlapati, Z. Feng, R.L. Prentice, S.M. Hanash, Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women, Cancer Prev. Res. (Phila). 5 (2012) 655–64. doi:10.1158/1940-6207.CAPR-11-0412.

[12] S. Werner, F. Krause, V. Rolny, M. Strobl, D. Morgenstern, C. Datz, H. Chen, H. Brenner, Evaluation of a 5-Marker Blood Test for Colorectal Cancer Early Detection in a Colorectal Cancer Screening Setting, Clin. Cancer Res. 22 (2016) 1725–1733. doi:10.1158/1078-0432.CCR-15-1268.

[13] S. Shao, B.A. Neely, T.-C. Kao, J. Eckhaus, J. Bourgeois, J. Brooks, E.E. Jones, R.R. Drake, K. Zhu, Proteomic Profiling of Serial Prediagnostic Serum Samples for Early Detection of Colon Cancer in the U.S. Military, Cancer Epidemiol. Biomarkers Prev. 26 (2017) 711–718. doi:10.1158/1055-9965.EPI-16-0732.

[14] J. Rho, J.J. Ladd, C.I. Li, J.D. Potter, Y. Zhang, D. Shelley, D. Shibata, D. Coppola, H. Yamada, H. Toyoda, T. Tada, T. Kumada, D.E. Brenner, S.M. Hanash, P.D. Lampe, Protein and glycomic plasma markers for early detection of adenoma and colon cancer, Gut. 67 (2018) 473–484. doi:10.1136/gutjnl-2016-312794.

[15] M. Bertuzzi, C. Marelli, R. Bagnati, A. Colombi, R. Fanelli, C. Saieva, M. Ceroti, B. Bendinelli, S. Caini, L. Airoldi, D. Palli, Plasma clusterin as a candidate pre-diagnosis marker of colorectal cancer risk in the Florence cohort of the European Prospective Investigation into Cancer and Nutrition: a pilot study, BMC Cancer. 15 (2015) 56. doi:10.1186/s12885-015-1058-7.

[16] T.F. Imperiale, D.F. Ransohoff, S.H. Itzkowitz, T.R. Levin, P. Lavin, G.P. Lidgard, D.A. Ahlquist, B.M. Berger, Multitarget stool DNA testing for colorectal-cancer screening, N. Engl. J. Med. 370 (2014) 1287–97. doi:10.1056/NEJMoa1311194.

[17] D.G. Redwood, E.D. Asay, I.D. Blake, P.E. Sacco, C.M. Christensen, F.D. Sacco, J.J. Tiesinga, M.E. Devens, S.R. Alberts, D.W. Mahoney, T.C. Yab, P.H. Foote, T.C. Smyrk, E.M. Provost, D.A. Ahlquist, Stool DNA Testing for Screening Detection of Colorectal Neoplasia in Alaska Native People, Mayo Clin. Proc. 91 (2016) 61–70. doi:10.1016/j.mayocp.2015.10.008.

[18] T.R. Church, M. Wandell, C. Lofton-Day, S.J. Mongin, M. Burger, S.R. Payne, E. Castaños-Vélez, B.A. Blumenstein, T. Rösch, N. Osborn, D. Snover, R.W. Day, D.F. Ransohoff, PRESEPT Clinical Study Steering Committee, Investigators and Study Team, Prospective evaluation of methylated *SEPT9* in plasma for detection of asymptomatic colorectal cancer, Gut. 63 (2014) 317–325. doi:10.1136/gutjnl-2012-304149.

[19] D.A. Johnson, R.L. Barclay, K. Mergener, G. Weiss, T. König, J. Beck, N.T. Potter, Plasma Septin9 versus Fecal Immunochemical Testing for Colorectal Cancer Screening: A Prospective Multicenter Study, PLoS One. 9 (2014) e98238. doi:10.1371/journal.pone.0098238.

[20] A. Adler, S. Geiger, A. Keil, H. Bias, P. Schatz, T. DeVos, J. Dhein, M. Zimmermann, R. Tauber, B. Wiedenmann, Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany, BMC Gastroenterol. 14 (2014) 183. doi:10.1186/1471-230X-14-183.

[21] D. Wu, G. Zhou, P. Jin, J. Zhu, S. Li, Q. Wu, G. Wang, J. Sheng, J. Wang, L. Song, X. Han, J. Qian, Detection of Colorectal Cancer Using a Simplified SEPT9 Gene Methylation Assay Is a Reliable Method for Opportunistic Screening, J. Mol. Diagnostics. 18 (2016) 535–545. doi:10.1016/j.jmoldx.2016.02.005.

Chapter 8. Afterword

A little over half-a-Century ago, the gFOBT was just a curiosity; fashioned into existence in an internist's office in a small corner of Ohio, in the USA. On publishing his findings in the *Journal of the American Medical Association* [1], little could the inventor, Dr David Greegor, foresee the impact his carboard apparatus would have on the prognosis of previvors whose CRC was screen-detected before any symptoms presented. Nowadays, the just-as-rudimentary kit greats thousands of Britons on their sixtieth birthday—soon to be fiftieth [2]—and each two years thereafter. Despite its undoubted benefits [3], however, the gFOBT is soon-to-be retired by the NHS and replaced with the newer generation FIT [4]. Nonetheless, the rationale remains: that premalignant and malignant tumours may leak occult blood detectable on stools before the onset of symptoms.

The current work was motivated on maximising the benefits of the BCSP through risk-stratified eligibility while also working toward an alternative blood test that may appeal to the 46% who avert conventional stool-based screening [5]. In working toward theses aims, it was demonstrated in Chapter 3 that 92% of CRCs diagnosed in England, Northern Ireland & Wales during 2002–11 were formally registered within one year of diagnosis—and that 99% were registered within six years [6]. Notably, these findings provide revised estimates on the comprehensiveness and reliability of UK cancer registries [7,8]. One would expect that, as the registration processes streamlines, the scientific community would benefit from these high-quality and data-rich registrations with less delay. Meanwhile, for those studies requiring data more immediately for their analyses, we have demonstrated that Hospital Episode Statistics in England tend to capture CRC events unregistered one year after diagnosis. This finding is particularly salient at a time of much fanfare for leveraging BigData to improve healthcare. Indeed, in late 2018, the UK government announced its intent and financial backing to 'use data, Artificial Intelligence and innovation to transform the prevention, early diagnosis and treatment of chronic diseases by 2030' [9]. Our study justifies the use of EHRs to ascertain events of CRC in such pursuits.

In concluding this thesis I have been reminded by UCL to state the impact the current work may have. In particular: is risk-stratified screening for CRC viable? And are asymptomatic CRCs detectable with a blood test? On the evidence presented the former would seem more likely than the latter—as we will now see.

After implementation of a population-based screening programme, attention naturally turns to the use of risk-stratified screening to redress the balance of benefits-to-

harms in favour of benefits [10]. The study of identical and non-identical twins demonstrated that the non-heritable component of susceptibility to CRC is double that heritable factors [11], and we speculated that the use of non-genetic factors recorded by EHRs would invite the same proportion of CRCs as current eligibility based on being 50–74-years old—with fewer screens overall. While it is true that this not the first risk-prediction model informed solely on routinely available non-genetic data [12–15], the use of these for risk reclassification has not yet been reported. Nevertheless, the results of Chapter 4 suggest that our NGRS would be less efficient than eligibility on age alone. A few silverlinings were that the NGRS would I) have correctly triaged adults as young as 33-years old for screening whom stand to benefit the most in years saved and II) could, at least in theory and pending further study, identify those of substantially higher risk who may benefit from colonoscopy over stool-based screening. We have also identified several ways in which the model can be refined; be it through longitudinal modelling of covariates, avoiding overdiagnosis in the elderly, inclusion of age into risk-predictions, and interpretation absolute risk scale. Having made these refinements, the performance of our NGRS could be evaluated in relation to existing models [15].

The second half of this thesis was committed towards a blood test able to screen for CRC. This was motivated on the outcomes of the BCSP during 2006–08—the first few years of the BCSP—for which just 54% underwent screening following their invitation [5] Indeed, uptake has decreased year-on-year since 2010 [16]. Low uptake undermines the potential benefits a population-based screening programme can provide. Would a blood test therefore be more acceptable to the general population? Preliminary evidence would suggest so [17,18].

In Chapter 5 our literature review of existing markers shortlisted the most able diagnostic markers: CEA, CYFRA21-1, FAP, TIMP-1, and VEGFA. We subsequently evaluated their ability to screen-detect CRCs using longitudinal preclinical samples drawn up to four years before diagnosis. CA125 was also evaluated in parallel using measurements of UKCTOCS multimodal screens. We show that CEA was elevated in up to one year in advance for 23% of future patients and in 14% between one and two years. These estimates, however, suggest it being inferior to the gFOBT used by the BCSP. This work, nevertheless, contributed to our knowledge of the performance of CEA as a function of lead time as well as its temporal variation in the years preceding CRC diagnosis [19]. Future work could use our data to devise a screening algorithm where referral to colonoscopy is triggered by temporal deviation from a stable CEA profile. The study also

demonstrated that neither CYFRA21-1, FAP, TIMP-1, nor VEGFA has sufficient sensitivity for deployment in population-based screening. The study reaffirms the need to study preclinical rather than diagnostic samples for future marker discovery work.

We thereafter set out in Chapter 6 to leverage our unique set of longitudinal preclinical samples in pursuit of novel screening markers. Very few others have undertook discovery work on preclinical samples [20–22]—and even less on longitudinal preclinical samples [23]. After substantially reducing the complexity of sera through the removal of highly abundant proteins and multi-dimensional fractionation, serum pools were submitted to mass analyses and the relative quantitation between pools was assessed. DCD and S100A8 were shortlisted on their Foldscore—a novel method of ranking proteins on their differential expression across multiple dimensions. Our internal validation on individual sera showed DCD and S100A8 showed only 10% positivity a year in advance of diagnosis. At the time of study, it was true that the even the most current mass spectrometers were able to analyse only a small proportion of all peptides injected [24]. The observed negative findings could thus be due to technical constraints and the inability of the technology used to sample the sub-proteome.

The experimental component of the thesis concluded with an exploration of all 248 possible multi-marker permutations of all protein measurements obtained. This was undertaken on rationale that other markers may supplement the shortcomings of CEA. All permutations were ranked by brute force and a five-marker panel of CA125, CEA, FAP, TIMP-1, and VEGFA ranked the highest but performed no better than CEA alone. Thus, it is difficult to justify the additional expenses of multi-marker panel testing.

All evidence considered; it is difficult to envision the FIT being replaced in the near future given that it performs as well as the genomic tests granted regulatory approval at a mere fraction of the cost [25,26]. For those with substantially high-risk and whom repeatedly avert undertaking the FIT, provision of the Epi proColon plasma m*SEPT9* test may be justified [17]. In the meanwhile, efforts to optimise the provision of CRC screening based on blood testing and/or flexible sigmoidoscopy could perhaps prove fruitful in the short-term. One would also expect an improvement in uptake as the FIT replaces the gFOBT [27,28]. Independently, initiatives based on behavioural psychology have delivered marginal gains to uptake in the past and could be worthwhile pursuing further [29–32]. Meanwhile, the UK National Screening Committee has commissioned a study on how best to allocate the fine NHS endoscopic capacity between their two CRC

screening programmes faced on stool occult blood testing and flexible sigmoidoscopy, respectively [2].

Overall, and despite our best efforts, no circulating proteins of those evaluated from the literature or our laboratory work have shown promise for detecting asymptomatic CRC. Irregardless, the publication of negative findings is an important contribution to evidence and could in this instance prevent retracing trodden steps [33]. Indeed, on the recurrent failures of screening markers, perhaps it should be questioned if the detection of a small, asymptomatic tumour via proteins leaked into circulation is theoretically achievable.

8.1 References

[1] D.H. Greegor, Diagnosis of large-bowel cancer in the asymptomatic patient, JAMA. 201 (1967) 943–5.

[2] UK National Screening Committee, Note of the meeting held on the 29 June 2018 in Edinburgh, Scotland, 2018.

[3] P. Hewitson, P. Glasziou, E. Watson, B. Towler, L. Irwig, Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update, Am J Gastroenterol. 103 (2008) 1541–1549. doi:10.1111/j.1572-0241.2008.01875.x.

[4] UK National Screening Committee, Note of the meeting held on the 19 November 2015 at Goodenough College, London, 2015.

[5] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut. 61 (2012) 1439–46. doi:10.1136/gutjnl-2011-300843.

[6] D.S. Thomas, A. Gentry-Maharaj, A. Ryan, E.-O. Fourkala, S. Apostolidou, M. Burnell, W. Alderton, J. Barnes, J.F. Timms, U. Menon, Colorectal cancer ascertainment through cancer registries, hospital episode statistics, and self-reporting compared to confirmation by clinician: A cohort study nested within the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), Cancer Epidemiol. 58 (2019) 167–174. doi:10.1016/J.CANEP.2018.11.011.

[7] H. Møller, S. Richards, N. Hanchett, S.P. Riaz, M. Lüchtenborg, L. Holmberg, D. Robinson, Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates, Br. J. Cancer. 105 (2011) 170–176. doi:10.1038/bjc.2011.168.

[8] A. Dregan, H. Møller, T. Murray-Thomas, M.C. Gulliford, Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study, Cancer Epidemiol. 36 (2012) 425–429. doi:10.1016/j.canep.2012.05.013.

[9] Department for Business; Energy & Industrial Strategy, Policy paper: The Grand Challenge missions, (2018). https://www.gov.uk/government/publications/industrial-strategy-the-grand-challenges (accessed March 1, 2019).

[10] D. Easton, N. Pashayan, P. Pharoah, Risk-stratified cancer screening, 2017.

[11] P. Lichtenstein, N. V Holm, P.K. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe, K. Hemminki, Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland, N Engl J Med. 343 (2000) 78–85. doi:10.1056/NEJM200007133430201.

[12] A.K. Win, R.J. Macinnis, J.L. Hopper, M.A. Jenkins, Risk prediction models for colorectal cancer: a review, Cancer Epidemiol. Biomarkers Prev. 21 (2012) 398–410. doi:10.1158/1055-9965.EPI-11-0771.

[13] G.K. Ma, U. Ladabaum, Personalizing Colorectal Cancer Screening: a Systematic Review of Models to Predict Risk of Colorectal Neoplasia, Clin. Gastroenterol. Hepatol. 12 (2014) 1624–1634.e1. doi:10.1016/j.cgh.2014.01.042.

[14] J.A. Usher-Smith, F.M. Walter, J.D. Emery, A.K. Win, S.J. Griffin, Risk Prediction Models for Colorectal Cancer: a Systematic Review, Cancer Prev. Res. 9 (2016) 13–26. doi:10.1158/1940-6207.CAPR-15-0274.

[15] J.A. Usher-Smith, A. Harshfield, C.L. Saunders, S.J. Sharp, J. Emery, F.M. Walter, K. Muir, S.J. Griffin, External validation of risk prediction models for incident colorectal cancer using UK Biobank, Br. J. Cancer. 118 (2018) 750–759. doi:10.1038/bjc.2017.463.

[16] Y. Hirst, S. Stoffel, G. Baio, L. McGregor, C. von Wagner, Uptake of the English Bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out, Eur. J. Cancer. 103 (2018) 267–273. doi:10.1016/j.ejca.2018.07.135.

[17] A. Adler, S. Geiger, A. Keil, H. Bias, P. Schatz, T. DeVos, J. Dhein, M. Zimmermann, R. Tauber, B. Wiedenmann, Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany, BMC Gastroenterol. 14 (2014) 183. doi:10.1186/1471-230X-14-183.

[18] E.G. Liles, G.D. Coronado, N. Perrin, A.H. Harte, R. Nungesser, N. Quigley, N.T. Potter, G. Weiss, T. Koenig, T. DeVos, Uptake of a colorectal cancer screening blood test is higher than of a fecal test offered in clinic: a randomized trial, Cancer Treat. Res. Commun. 10 (2017) 27–31. doi:10.1016/J.CTARC.2016.12.004.

[19] D.S. Thomas, E.-O. Fourkala, S. Apostolidou, R. Gunu, A. Ryan, I. Jacobs, U. Menon, W. Alderton, A. Gentry-Maharaj, J.F. Timms, Evaluation of serum CEA, CYFRA21-1 and CA125 for the early detection of colorectal cancer using longitudinal preclinical samples, Br. J. Cancer. 113 (2015) 268–74. doi:10.1038/bjc.2015.202.

[20] J.J. Ladd, T. Busald, M.M. Johnson, Q. Zhang, S.J. Pitteri, H. Wang, D.E. Brenner, P.D. Lampe, R. Kucherlapati, Z. Feng, R.L. Prentice, S.M. Hanash, Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women, Cancer Prev. Res. (Phila). 5 (2012) 655–64. doi:10.1158/1940-6207.CAPR-11-0412.

[21] M. Bertuzzi, C. Marelli, R. Bagnati, A. Colombi, R. Fanelli, C. Saieva, M. Ceroti, B. Bendinelli, S. Caini, L. Airoldi, D. Palli, Plasma clusterin as a candidate pre-diagnosis marker of colorectal cancer risk in the Florence cohort of the European Prospective Investigation into Cancer and Nutrition: a pilot study, BMC Cancer. 15 (2015) 56. doi:10.1186/s12885-015-1058-7.

[22] J. Rho, J.J. Ladd, C.I. Li, J.D. Potter, Y. Zhang, D. Shelley, D. Shibata, D. Coppola, H. Yamada, H. Toyoda, T. Tada, T. Kumada, D.E. Brenner, S.M. Hanash, P.D. Lampe, Protein and glycomic plasma markers for early detection of adenoma and colon cancer, Gut. 67 (2018) 473–484. doi:10.1136/gutjnl-2016-312794.

[23] S. Shao, B.A. Neely, T.-C. Kao, J. Eckhaus, J. Bourgeois, J. Brooks, E.E. Jones, R.R. Drake, K. Zhu, Proteomic Profiling of Serial Prediagnostic Serum Samples for Early Detection of Colon Cancer in the U.S. Military, Cancer Epidemiol. Biomarkers Prev. 26 (2017) 711–718. doi:10.1158/1055-9965.EPI-16-0732.

[24] A. Michalski, J. Cox, M. Mann, More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS, J. Proteome Res. 10 (2011) 1785–93. doi:10.1021/pr101060v.

[25] H. Brenner, S. Werner, H. Chen, Correspondence: Multitarget Stool DNA Testing for Colorectal-Cancer Screening, N. Engl. J. Med. 371 (2014) 184–188. doi:10.1056/NEJMc1405215.

[26] D.A. Johnson, R.L. Barclay, K. Mergener, G. Weiss, T. König, J. Beck, N.T. Potter, Plasma Septin9 versus Fecal Immunochemical Testing for Colorectal Cancer Screening: A Prospective Multicenter Study, PLoS One. 9 (2014) e98238. doi:10.1371/journal.pone.0098238.

[27] J. Digby, P.J. McDonald, J.A. Strachan, G. Libby, R.J. Steele, C.G. Fraser, Use of a faecal immunochemical test narrows current gaps in uptake for sex, age and deprivation in a bowel cancer screening programme, J. Med. Screen. 20 (2013) 80–85. doi:10.1177/0969141313497197.

[28] S. Moss, C. Mathews, T.J. Day, S. Smith, H.E. Seaman, J. Snowball, S.P. Halloran, Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England, Gut. 66 (2017) 1631–1644. doi:10.1136/gutjnl-2015-310691.

[29] J. Wardle, C. von Wagner, I. Kralj-Hans, S.P. Halloran, S.G. Smith, L.M. McGregor, G. Vart, R. Howe, J. Snowball, G. Handley, R.F. Logan, S. Rainbow, S. Smith, M.C. Thomas, N. Counsell, S. Morris, S.W. Duffy, A. Hackshaw, S. Moss, W. Atkin, R. Raine, Effects of evidence-based strategies to reduce the socioeconomic gradient of uptake in the English NHS Bowel Cancer Screening Programme (ASCEND): four cluster-randomised controlled trials, Lancet (London, England). 387 (2016) 751–9. doi:10.1016/S0140-6736(15)01154-X.

[30] R. Raine, S.M. Moss, C. von Wagner, W. Atkin, I.K. Hans, R. Howe, F. Solmi, S. Morris, N. Counsell, A. Hackshaw, S. Halloran, G. Handley, R.F. Logan, S. Rainbow, S. Smith, J. Snowball, H. Seaman, M. Thomas, S.G. Smith, L.M. McGregor, G. Vart, J. Wardle, S.W. Duffy, A national cluster-randomised controlled trial to examine the effect of enhanced reminders on the socioeconomic gradient in uptake in bowel cancer screening, Br. J. Cancer. 115 (2016) 1479–1486. doi:10.1038/bjc.2016.365.

[31] P. Hewitson, A.M. Ward, C. Heneghan, S.P. Halloran, D. Mant, Primary care endorsement letter and a patient leaflet to improve participation in colorectal cancer screening: results of a factorial randomised trial, Br. J. Cancer. 105 (2011) 475–80. doi:10.1038/bjc.2011.255.

[32] J. Shankleman, N.J. Massat, L. Khagram, S. Ariyanayagam, A. Garner, S. Khatoon, S. Rainbow, S. Rangrez, Z. Colorado, W. Hu, D. Parmar, S.W. Duffy, Evaluation of a service intervention to improve awareness and uptake of bowel cancer screening in ethnically-diverse areas, Br. J. Cancer. 111 (2014) 1440–7. doi:10.1038/bjc.2014.363.

[33] K. Dickersin, I. Chalmers, Recognizing, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the WHO, J. R. Soc. Med. 104 (2011) 532–8. doi:10.1258/jrsm.2011.11k042.

Supplementary Appendices

Supplementary Method M1

Specificity & NPV. Specificities and NPVs were estimated relative to the expected number of cases derived from an age-standardised rate of 57.2 cases per 100,000 person years [1] applied to the total years of follow-up for subsets of all 202,365 consenting women who 1) lived in the UK (CR analysis), 2) had died before the latest DC update (DC analysis), 3) lived in England (HES analysis), and 4) who returned their FUQ before 24 May 2011 (SR analysis). Total years of follow-up was the sum of all years followed-up from randomisation to latest update (see *Electronic health records*), or date of death if before. Column totals of the confusion matrix were inferred from the expected number of cases (TP+FN) or sample size minus expected cases (FP+TN). Row totals of the confusion matrix were inferred from the number of notifications (TP+FP) or sample size minus number of notifications (FN+TN). Using the sensitivity estimates (Table 3—3 or Supplementary Table M1 for non-rounded, crude estimates), the number of TPs can be calculated as 1) sensitivity * (TP+FN). Individual cells (i.e. FN, TP & TN) were then deduced from subtraction of TP from column/row totals. The specificity (TN/(FP+TN)), NPV (TN/(FN+TN)), and 95% confidence intervals were then computed.

Supplementary Table M1: Crude Sensitivities and Positive Predictive Values.

Dataset	Crude sensitivity	Crude PPV
CR[1]	0.9226069	0.9476987
DC	0.9702970	0.9800000
HES	0.8236776	0.9369628
SR	0.9055794	0.6895425

[1] 1–9 years curation (median 4.1, IQR 3.2). Abbreviations: CIs, confidence intervals; PPV, positive predictive value; CR, cancer registration; DC, death certificate; HES, Hospital Episode Statistics; SR, self-reporting.

Proofs: Cancer registrations. There were 814 (+) cancer registrations received by latest registry update and 850 expected cancers:

TP + FN = 850 expected cancers.

FP + TN = 201,515 (202,365 consenting women followed-up minus 850 expected cancers).
TP + FP = 814 (+) cancer registrations received by latest update.
FN + TN = 201,551 (202,365 consenting women followed-up minus 814 (+) cancer registrations).

Since sensitivity of cancer registrations relative to clinical confirmations was 0.9226069, the number of TPs can be calculated as sensitivity * (TP + FN), or 0.9226069 * 850, which equals 784. FN is therefore 66 (850 minus 784), FP is 30 (814 minus 784), and TN is 201,485 (201,515 minus 30 or 201,551 minus 66).

Confusion Matrix According to Cancer Registrations Received and Expected Number of Cases.

	Expected +	Expected –	Total
Notification +	784	30	814
Notification –	66	201,485	201,551
Total	850	201,515	202,365

Proofs: Death registrations. There were 233 (+) death registrations received by latest registry update and 19 expected cancers:

TP + FN = 19 expected cancers.
FP + TN = 7,183 (7,202 women deceased minus 19 expected cancers).
TP + FP = 233 (+) death registrations received by latest update.
FN + TN = 6,969 (2,202 deceased women minus 233 (+) death registrations).

Since sensitivity of death registrations relative to clinical confirmations was 0.9702970, the number of TPs can be calculated as sensitivity * (TP + FN), or 0.9702970 * 19, which equals 18. FN is therefore 1 (19 minus 18), FP is 215 (233 minus 18), and TN is 6,968 (7,183 minus 215 or 6,969 minus 1).

Confusion Matrix According to Death Registrations Received and Expected Number of Cases.

	Expected +	Expected –	Total
Notification +	18	215	233
Notification –	1	6,968	6,969
Total	19	7,183	7,202

Example: Hospital Episode Statistics. There were 625 (+) Hospital Episode Statistics received by latest update and 616 expected cancers:

TP + FN = 616 expected cancers.
FP + TN = 157,223 (157,839 women followed-up minus 616 expected cancers).
TP + FP = 625 (+) Hospital Episode Statistics received by latest update.
FN + TN = 157,214 (157,839 women followed-up minus 625 (+) Hospital Episode Statistics).

Since sensitivity of Hospital Episode Statistics relative to clinical confirmations was 0.8236776, the number of TPs can be calculated as sensitivity * (TP + FN), or 0.8236776 * 616, which equals 507. FN is therefore 109 (616 minus 507), FP is 118 (625 minus 507), and TN is 157,105 (157,223 minus 118 or 157,214 minus 109).

Confusion Matrix According to Hospital Episode Statistics Received and Expected Number of Cases.

	Expected +	**Expected –**	**Total**
Notification +	507	118	625
Notification –	109	157,105	157,214
Total	616	157,223	157,839

Example: Self-reporting. There were 400 (+) self-reportings received and 321 expected cancers:

TP + FN = 321 expected cancers.
FP + TN = 143,992 (144,313 women followed-up minus 321 expected cancers).
TP + FP = 400 (+) self-reportings received.
FN + TN = 143,913 (144,313 women followed-up minus 400 (+) self-reportings).

Since sensitivity of Hospital Episode Statistics relative to clinical confirmations was 0.9055794, the number of TPs can be calculated as sensitivity * (TP + FN), or 0.9055794 * 321, which equals 291. FN is therefore 30 (321 minus 291), FP is 109 (400 minus 291), and TN is 143,883 (143,992 minus 109 or 143,913 minus 30).

Confusion Matrix According to Self-reportings Received and Expected Number of Cases.			
	Expected +	**Expected –**	**Total**
Notification +	291	109	400
Notification –	30	143,883	143,913
Total	321	143,992	144,313

[1] Cancer Research UK, Bowel cancer incidence statistics [UK, 2015, ICD-10 C18–C20], (2018). http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence (accessed February 28, 2018).

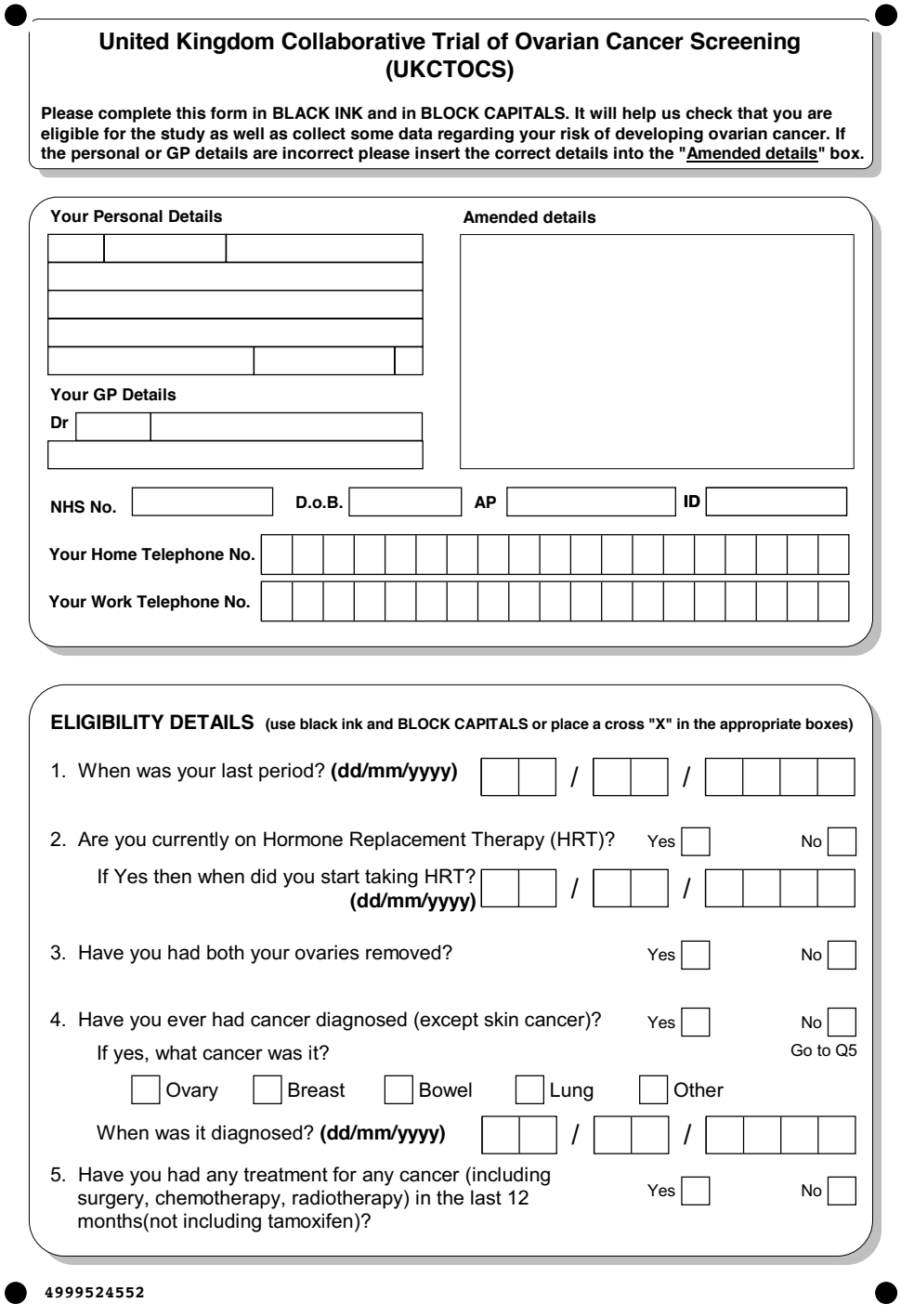
Supplementary Table S1: Proteins Targeted by Immunodepletion Kits.	
Proteome Purify 12 (R&D Systems)	**ProteoPrep 20 (Sigma-Aldrich)**
α1-Acid Glycoprotein	α1-Acid Glycoprotein
α1-Antitrypsin	α1-Antitrypsin
α2-Macroglobulin	α2-Macroglobulin
Albumin	Albumin
Apolipoprotein A-I	Apolipoprotein A-I
Apolipoprotein A-II	Apolipoprotein A-II
Fibrinogen	Fibrinogen
Haptoglobin	Haptoglobin
IgA	IgA
IgG	IgG
IgM	IgM
Transferrin	Transferrin
	IgD
	Ceruloplasmin
	Apolipoprotein B
	Complement C1q subunit A
	Complement C1q subunit B
	Complement C3
	Complement C4-A
	Complement C4-B
	Plasminogen
	Transthyretin

Supplementary Table S2: Excluded Contaminants.

cRAP	Max Plank Institute
Keratin, type I cuticular Ha1	Dermokine beta-1
Keratin, type I cuticular Ha2	Dermokine
Keratin, type I cuticular Ha3-I	Keratin, type II cuticular Hb6
Keratin, type I cuticular Ha3-II	Keratin, type I cuticular HA3-I
Keratin, type I cuticular Ha4	Keratin, type I cuticular HA4
Keratin, type I cuticular Ha5	Keratin, type I cuticular HA6
Keratin, type I cuticular Ha6	Keratin, type I cuticular HA7
Keratin, type I cuticular Ha7	Keratin, type I cuticular HA8
Keratin, type I cuticular Ha8	Keratin, type II cytoskeletal 75
Keratin, type I cytoskeletal 10	Keratin, type I cytoskeletal 14
Keratin, type I cytoskeletal 9	Keratin, type II cytoskeletal 6A
Keratin, type II cuticular Hb1	Keratin, type II cytoskeletal 6B
Keratin, type II cuticular Hb2	Keratin, type II cytoskeletal 1
Keratin, type II cuticular Hb3	Keratin, type II cytoskeletal 8
Keratin, type II cuticular Hb4	Keratin, type I cytoskeletal 19
Keratin, type II cuticular Hb5	Keratin, type II cytoskeletal 7
Keratin, type II cuticular Hb6	Keratin, type I cytoskeletal 16
Keratin, type II cytoskeletal 1	Keratin, type II cytoskeletal 3
Keratin, type II cytoskeletal 2 epidermal	Keratin, type I cytoskeletal 10
Salivary alpha-amylase precursor	Keratin, type I cytoskeletal 13
	Keratin, type II cytoskeletal 5
	Keratin, type I cytoskeletal 15
	Keratin 4
	Filaggrin
	Keratin, type I cytoskeletal 9
	Keratin, type I cytoskeletal 20
	Keratin, type II cytoskeletal 2 epidermal
	Keratin, type II cytoskeletal 6C
	Keratin-associated protein 10-12
	Keratin, type II cuticular Hb3
	Keratin, type II cuticular Hb5
	Keratin, type II cytoskeletal 2 oral
	Keratin, type I cytoskeletal 17
	Keratin, type I cuticular HA3-II
	Keratin, type I cuticular HA2
	Keratin, type II cuticular Hb1
	Keratin, type II cytoskeletal 72
	Keratin, type I cuticular HA1
	Keratin, type I cytoskeletal 24
	Keratin-73
	KRT7 protein

	Keratin-associated protein 19-5
	Keratin, type II cytoskeletal 71
	Filaggrin-2
	Keratin, type II cytoskeletal 79
	Type I hair keratin KA36
	Type I hair keratin KA35
	Keratin, hair, basic, 4
	Keratin, type II cytoskeletal 80
	Keratin, hair, basic, 3
	Type I hair keratin 3A
	Keratin, type II cytoskeletal 74
	Keratin-78
	Keratin 25D
	Keratin, type I cytoskeletal 27
	Keratin, type I cytoskeletal 26
	Keratin, type I cytoskeletal 25
	Keratin 77
	Hornerin
	Type I hair keratin 4
	Keratin-like protein KRT222
	Keratin, type II cytoskeletal 78
	Keratin, type I cuticular HA5
	Keratin, type I cytoskeletal 12
	Keratin-associated protein 4-6
	Keratin-associated protein 4-1
	Keratin-associated protein 4-9
	Keratin-associated protein 4-3
	Keratin-associated protein 3-1
	Keratin-associated protein 2-4
	Keratin, type I cytoskeletal 23
	Keratin-8-like protein 1
	Keratin, type II cuticular Hb4
	Keratin, type II cuticular Hb2
	Type I hair keratin 1

Contaminants compiled in the common Repository of Adventitious Proteins (https://www .thegpm.org/cRAP/) and Max Plank Institute (http://www.maxquant.org/downloads.htm). Abbreviations: cRAP, common Repository of Adventitious Proteins.



United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

Please complete this form in BLACK INK and in BLOCK CAPITALS. It will help us check that you are eligible for the study as well as collect some data regarding your risk of developing ovarian cancer. If the personal or GP details are incorrect please insert the correct details into the "Amended details" box.

Your Personal Details

Amended details

Your GP Details

Dr

NHS No. D.o.B. AP ID

Your Home Telephone No.

Your Work Telephone No.

ELIGIBILITY DETAILS (use black ink and BLOCK CAPITALS or place a cross "X" in the appropriate boxes)

1. When was your last period? **(dd/mm/yyyy)** / /

2. Are you currently on Hormone Replacement Therapy (HRT)? Yes No

   If Yes then when did you start taking HRT? **(dd/mm/yyyy)** / /

3. Have you had both your ovaries removed? Yes No

4. Have you ever had cancer diagnosed (except skin cancer)? Yes No Go to Q5

   If yes, what cancer was it?

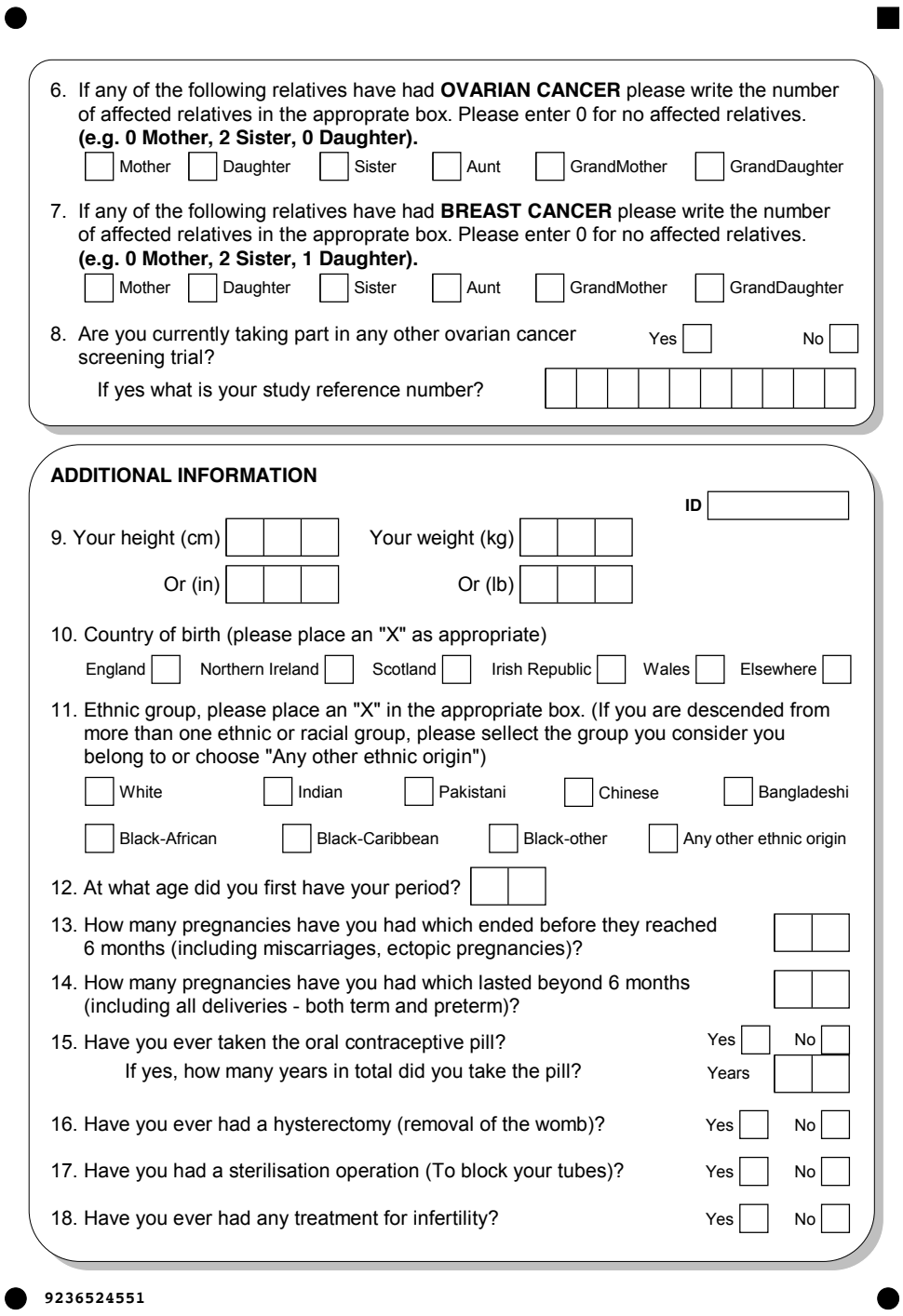
   Ovary Breast Bowel Lung Other

   When was it diagnosed? **(dd/mm/yyyy)** / /

5. Have you had any treatment for any cancer (including surgery, chemotherapy, radiotherapy) in the last 12 months(not including tamoxifen)? Yes No

4999524552

Supplementary Figure S1: Page One of UKCTOCS Recruitment Questionnaire.



Supplementary Figure S1: Page Two of UKCTOCS Recruitment Questionnaire.

0175097515 1 of 4

Volunteer Ref:

U K C T O C S

Private and Confidential

United Kingdom Collaborative Trial of Ovarian Cancer Screening

For the information that you supply to be useful to our ovarian cancer screening study, it is important that you complete all of the questions in "bold"

Please use a BLACK pen and place a cross INSIDE the box ☒

Please return both sheets using the Free Post envelope supplied

Follow up questionnaire

We would be very grateful if you could answer the following questions. If you are not sure about exact details/dates an approximate answer is better than none. If there are any relevant details you wish to include, please use an additional sheet. Please use a BLACK biro or ink pen.

General questions about you

What qualification(s) do you have from school, college or the equivalent?

(please place a cross inside the most appropriate box(es))

☐ "O" level or equivalent
☐ "A" level or equivalent
☐ Clerical or commercial qualification (e.g. secretarial, hairdressing etc)
☐ None of these
☐ Nursing or teaching
☐ College/ university degree (or equivalent)

Approximately how much alcohol on average do you drink each week? (One drink = a glass of wine, half a pint of lager or cider, a measure of spirits). Average number of drinks of alcohol each week :

☐ None ☐ Less than 1 ☐ 1-3 ☐ 4-6 ☐ 7-10 ☐ 11-15 ☐ 16-20 ☐ 21+

Have you ever been a smoker? ☐ Yes ☐ No

If you answered yes to the above please answer the following questions:

How many years in total have you smoked for? ☐☐

During those years how many cigarettes on average did you smoke per day? ☐☐

What was your skirt size when you were in your early twenties?

☐ 6 ☐ 8 ☐ 10 ☐ 12 ☐ 14 ☐ 16 ☐ 18 ☐ 20 ☐ 22 ☐ 24 ☐ 26 ☐ 28 ☐ 30

What is your skirt size now?

☐ 6 ☐ 8 ☐ 10 ☐ 12 ☐ 14 ☐ 16 ☐ 18 ☐ 20 ☐ 22 ☐ 24 ☐ 26 ☐ 28 ☐ 30

Are you currently taking HRT? ☐ No ☐ Yes

Have you used any of the following to relieve menopausal symptoms? ☐ Yes ☐ No

☐ Herbal remedies e.g. Black cohosh
☐ Phytoestrogens or soy products
☐ Vitamins e.g. Menopace, vitamin E
☐ Other medical treatments e.g. Venlafaxine, Megace
☐ Homeopathic remedies
☐ Aromatherapy, reflexology or acupuncture
☐ Life style changes e.g. relaxation, exercise

We are interested to know more about how women deal with the menopause. If you are 50-60 years old would you be willing to complete a survey? ☐ Yes ☐ No

Supplementary Figure S2: Page One of UKCTOCS Follow-up Questionnaire I.

6194097512

2 of 4

Questions about your outlook on life

Using the scale below, please indicate the extent to which you agree with each item:

1. The future seems to me to be hopeful, and I believe that things are changing for the better.

☐ Absolutely agree ☐ Somewhat agree ☐ Cannot say ☐ Somewhat disagree ☐ Absolutely disagree

2. I feel that it is possible to reach the goals I would like to strive for.

☐ Absolutely agree ☐ Somewhat agree ☐ Cannot say ☐ Somewhat disagree ☐ Absolutely disagree

Questions about your Health

Do you have/are you being treated for any of the following conditions (multiple boxes can be crossed)?

☐ High blood pressure ☐ Diabetes ☐ Stroke

☐ Heart disease e.g. heart attack, angina ☐ Rheumatoid arthritis ☐ Osteoporosis

☐ High blood cholesterol ☐ Osteoarthritis

☐ I have not been treated for any of the above conditions

Since joining UKCTOCS have you had any of the following operations? ☐ Yes ☐ No

If "yes" please enter the details (multiple boxes can be filled; please fill even if you had them as part of the trial)

Operation to look at your ovaries -either by incision or keyhole (laparoscopy)

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

Name of Consultant: _____

Hysterectomy / Removal of womb

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

Name of Consultant: _____

Removal of ovaries (please tick one box) ☐ Right Ovary ☐ Left Ovary ☐ Both ovaries

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

Name of Consultant: _____

Hysteroscopy / D&C / Scrape of your womb / Operation to look at your womb

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

Name of Consultant: _____

Were there any complications resulting from any of the above procedures? ☐ Yes ☐ No

If yes, please give details using an extra sheet of paper if necessary.

_____

_____

Supplementary Figure S2: Page Two of UKCTOCS Follow-up Questionnaire I

7093097511

3 of 4

Breast biopsy or other surgery involving the breast

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

_____

Name of Consultant: _____

Any other operation of any kind. Please describe:

Since joining UKCTOCS **have you been diagnosed with any cancer? Please tell us about this(ese) cancer(s)**

☐ Ovarian cancer	☐ Bowel/colorectal cancer	☐ Lung cancer
☐ Breast cancer	☐ Gastric/stomach cancer	☐ Vulval/vaginal cancer
☐ Cervical cancer	☐ Pancreatic cancer	☐ BCC/rodent/skin cancer
☐ Endometrial/uterus/womb cancer	☐ Kidney cancer	☐ Other cancer

☐ I have not been diagnosed with any cancers

Type of cancer: _____

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

_____

Name of Consultant: _____

Type of cancer: _____

Year of operation: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

_____

Name of Consultant: _____

Once you have completed this questionnaire please sign and date below, then return it in the Free Post envelope. Thanks you.

Name (please print): Date: ☐☐ / ☐☐ / ☐☐☐☐

Signature: (dd/mm/yyyy)

QUOX 1 ☐ 2 ☐ 3 ☐ **Office use only** **QU Consent** Yes ☐ No ☐

Supplementary Figure S2: Page Three of UKCTOCS Follow-up Questionnaire I

The following sections should ONLY be completed by those volunteers who are in the CONTROL group

Since joining UKCTOCS have you had an ultrasound scan of your ovaries? ☐ Yes ☐ No

If "yes" why was this performed?

☐ GP request

☐ Hospital Doctor request

☐ Your own request

☐ Other reason _____

Year of scan: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

_____

Name of Consultant: _____

Since joining UKCTOCS have you had a blood test for CA125? ☐ Yes ☐ No

(CA125 is a substance, which is released at higher levels into the blood in women with ovarian cancer. The test is carried out if doctors suspect that a woman may have ovarian cancer)

If "yes" why was this performed?

☐ GP request

☐ Hospital Doctor request

☐ Your own request

☐ Other reason _____

Year of CA125 test: _____ Hospital No.: _____

Hospital at which operation took place: _____

_____

_____

Name of Consultant: _____

Thank you for taking the time to complete this questionnaire. The information that you supply is of great importance to the success of the trial.

Regards

UKCTOCS Team

Supplementary Figure S2: Page Four of UKCTOCS Follow-up Questionnaire I

UKCTOCS

United Kingdom Collaborative Trial of Ovarian Cancer Screening

Second Follow-up Health Questionnaire

Thank you for taking part in UKCTOCS. Please answer every question as fully as you can.

If you are not sure of the exact dates/ details, an approximate answer is better than none.

Please use a **BLACK** pen and where appropriate place a cross [X] inside the box for your answer. Any information that you provide will be treated in the strictest of confidence and will only be used for research purposes. We may contact you if anything is unclear.

We appreciate your time and effort in completing this questionnaire. If possible, please return the questionnaire within 21 days of receipt.

SECTION I: YOUR CONTACT DETAILS

1. **Do you have an email address that we can use to contact you in the future?** Yes ☐ No ☐
   If yes, please provide below. Email address:

2. **If you have received this questionnaire as redirected mail, please enter your current contact details below so that we can update our records:**
   Address:
   Town: County: Postcode:

If your contact details change in the future, please do provide us with your new postal or email address. We will email you the results of the trial when analysis is completed in 2015 if you have an email address.

SECTION II: GENERAL INFORMATION THAT WE DID NOT ASK PREVIOUSLY

1. **Marital Status** (please tick only one box)
   (a) Current:
   ☐ Single ☐ Married ☐ Partnered ☐ Divorced ☐ Separated ☐ Widowed
   (b) Before trial started in 2000
   ☐ Single ☐ Married ☐ Partnered ☐ Divorced ☐ Separated ☐ Widowed

2. **What was your primary occupation?** (Please enter the role in which you have been engaged the longest)

Supplementary Figure S3: Page One of UKCTOCS Follow-up Questionnaire II

3. **Please measure and enter the following measurements:**

(a) Your height: ☐ feet ☐☐ inches (b) Your weight: ☐☐ stones ☐☐ lbs

(c) Your waist: ☐☐ inches (d) Your hips: ☐☐ inches

SECTION III: GYNAECOLOGICAL MEDICAL HISTORY – SURGICAL AND INVESTIGATIVE PROCEDURES

Have you had any of the following procedures since joining UKCTOCS?
If yes, please enter the details below. Multiple boxes can be filled; please fill even if you had the procedures as part of the trial.

1. **Operation to look at/remove your ovaries/fallopian tubes** Yes ☐ No ☐
2. **Operation to look into your womb (hysteroscopy), endometrial biopsy (removal of some tissue from your womb) or D & C (scrape of your womb)** Yes ☐ No ☐
3. **Operation to remove your womb (hysterectomy)** Yes ☐ No ☐
4. **Operation to remove part or whole of your breast(s) (e.g. lumpectomy/ wide local excision/mastectomy)** Yes ☐ No ☐

   Any other gynaecological surgery Yes ☐ No ☐
5. **If you have ticked 'Yes' to any of the above, please provide details**

a) Type of operation:

Date of surgery: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

b) Type of operation:

Date of surgery: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

SECTION IV: YOUR MEDICAL HISTORY

1. **Since joining UKCTOCS have you been diagnosed with any cancer?** *(multiple boxes can be filled)*

☐ **I have not been diagnosed with any cancers**

☐ Ovarian cancer	☐ Bowel/colorectal cancer	☐ Lymphoma
☐ Breast cancer	☐ Gastric/stomach cancer	☐ Leukaemia
☐ Cervical cancer	☐ Pancreatic cancer	☐ BCC/rodent/skin cancer
☐ Endometrial/womb cancer	☐ Kidney cancer	☐ Melanoma
☐ Vulval/vaginal cancer	☐ Bladder cancer	☐ Other cancer (please specify below)

2. **If you have ticked any of the above cancers, please provide details of the main treating hospital where you had your surgery or chemotherapy:**

a) Type of cancer: Date of diagnosis: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

Supplementary Figure S3: Page Two of UKCTOCS Follow-up Questionnaire II

b) Type of cancer: Date of diagnosis: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

3. **Do you have/are being treated for any of the following conditions?** *(multiple boxes can be filled)*

☐ **I do not have any of these conditions**

☐ High blood pressure
☐ Heart disease (e.g. heart attack, angina)
☐ High blood cholesterol
☐ Stroke
☐ Diabetes
☐ Rheumatoid arthritis
☐ Osteoarthritis
☐ Osteoporosis
☐ Thyroid disease
☐ Kidney disease
☐ Inflammatory bowel disease
☐ Liver disease
☐ Chronic Obstructive Pulmonary Disease

4. **If you have ticked any of the above conditions, please provide details of the main treating hospital:**

a) Condition: Date of diagnosis: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

b) Condition: Date of diagnosis: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

Hospital Name:

Hospital Number: Name of Consultant:

5. **Have you ever taken any of the following medications?**

Tamoxifen Yes ☐ No ☐ Start date ☐☐ / ☐☐ / ☐☐ Stop date ☐☐ / ☐☐ / ☐☐
Statins Yes ☐ No ☐ Start date ☐☐ / ☐☐ / ☐☐ Stop date ☐☐ / ☐☐ / ☐☐
Low-dose aspirin Yes ☐ No ☐ Start date ☐☐ / ☐☐ / ☐☐ Stop date ☐☐ / ☐☐ / ☐☐

SECTION V: ADDITIONAL OVARIAN CANCER SCREENING

1. **Have you had any screening for ovarian cancer OUTSIDE UKCTOCS?** Yes ☐ No ☐

2. **If yes, when was it done and where was this performed?** Date: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

☐ NHS Hospital Hospital name:
☐ GP surgery Practice name:
☐ UKCTOCS Trial Centre Name of Centre:
☐ Private Clinic Details of private clinic:

3. **Why was it done?**

☐ Symptoms ☐ Screening ☐ Your request ☐ Other reason (please specify)

Supplementary Figure S3: Page Three of UKCTOCS Follow-up Questionnaire II

4. **If yes, was it abnormal?** Yes ☐ No ☐
Did it result in additional tests? Yes ☐ No ☐
Did it result in surgery? Yes ☐ No ☐
Did it lead to a cancer diagnosis? Yes ☐ No ☐

5. **Since joining UKCTOCS have you had an ultrasound scan?** Yes ☐ No ☐

6. **If yes, when was it done and where was this performed?** Date: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

☐ NHS Hospital	Hospital name:
☐ GP surgery	Practice name:
☐ UKCTOCS Trial Centre	Name of Centre:
☐ Private Clinic	Details of private clinic:

7. **Why was it done?**

☐ Symptoms ☐ Screening ☐ Your request ☐ Other reason (please specify)

8. **If yes, was it abnormal?** Yes ☐ No ☐
Did it result in additional tests? Yes ☐ No ☐
Did it result in surgery? Yes ☐ No ☐
Did it lead to a cancer diagnosis? Yes ☐ No ☐

SECTION VI: YOUR FAMILY HISTORY OF CANCER

1. **If any of the following relatives have had OVARIAN CANCER please write the number of affected relatives in the appropriate box. Please enter 0 for no affected relatives (e.g. 0 Mother, 2 Sister, 0 Daughter)**

☐ Mother	☐ Daughter	☐ Paternal Aunt	☐ Paternal Grandmother
☐ Sister	☐ Granddaughter	☐ Maternal Aunt	☐ Maternal Grandmother

2. **If any of the following relatives have had BREAST CANCER please write the number of affected relatives in the appropriate box. Please enter 0 for no affected relatives (e.g. 0 Mother, 2 Sister, 1 Daughter)**

☐ Mother	☐ Daughter	☐ Paternal Aunt	☐ Paternal Grandmother
☐ Sister	☐ Granddaughter	☐ Maternal Aunt	☐ Maternal Grandmother

3. **If any of the following relatives have had ENDOMETRIAL CANCER please write the number of affected relatives in the appropriate box. Please enter 0 for no affected relatives (e.g. 0 Mother, 2 Sister, 1 Daughter)**

☐ Mother	☐ Daughter	☐ Paternal Aunt	☐ Paternal Grandmother
☐ Sister	☐ Granddaughter	☐ Maternal Aunt	☐ Maternal Grandmother

Please enter the date you completed this questionnaire: ☐☐ / ☐☐ / ☐☐ (dd/mm/yy)

On completion, please check to make sure you have answered **all** the questions. Please return the questionnaire to us in the **FREEPOST** envelope provided.

Thank you for taking the time to complete this questionnaire. The information that you supply is of great importance to the success of the trial.

Supplementary Figure S3: Page Four of UKCTOCS Follow-up Questionnaire II

United Kingdom Collaborative Trial of Ovarian Cancer Screening

UKCTOCS Coordinating Centre
Gynaecological Cancer Research Centre
Institute for Women's Health, UCL
Maple House, 149 Tottenham Court Road
London, W1T 7DN
Fax: 0203 447 2129

Private and Confidential
«ConsultantTitle» «ConsultantFirstName» «ConsultantSurname»
«ConsultantSpeciality»
«HospAddress1»
«HospAddress2»
«HospTown»
«HospPostCode»

Dear «ConsultantTitle» «ConsultantSurname»,

RE: «V_FirstNames» «V_Surname»
DOB: «V_DateOfBirth»
Address : «V_Address1», «V_Address2», «V_Town», «V_PostCode»
NHS number: «V_NHSNumber»
UKCTOCS reference number: «O_Volunteeref»

Mrs «V_Surname» is a participant in the UKCTOCS ovarian cancer screening trial, which is a major national MRC, CRUK, and DoH funded randomised controlled trial involving 202,638 women in the UK. According to the notifications we have received, Mrs «V_Surname» has been diagnosed with COLORECTAL CANCER. In order to fully utilise the serum samples that Mrs «V_Surname» has donated to cancer research we kindly ask you to complete the questions below, which relate to the histology at diagnosis and any treatment the patient may have received.

Please could you fill in the following questionnaire and send us a copy of the histology if available. A free post envelope is enclosed. I have enclosed a copy of Mrs «V_Surname»'s consent form to take part in the UKCTOCS which provides permission for access to her medical records.

If Mrs «V_Surname» was not under your care, it would be appreciated if you provide the contact details of the consultant responsible for Mrs «V_Surname»'s cancer treatment.

Name of consultant:……..

Address:……

……

Yours sincerely,

Professor Usha Menon

Name of patient: «V_FirstNames» «V_Surname» **DOB: «V_DateOfBirth»**
UKCTOCS reference number: «V_VolunteerRef» **NHS number: «V_NHSNumber»**

Supplementary Figure S4: Page One of Colorectal cancer questionnaire (CRCQ) sent to all treating clinicians requesting confirmation and histology reports.

United Kingdom Collaborative Trial of Ovarian Cancer Screening

Date of Diagnosis: _ _/ _ _/ _ _ _ _ (DD/MM/YYYY)

Primary Tumour Site:

☐ Ascending Colon ☐ Transverse Colon ☐ Descending Colon

☐ Sigmoid Colon ☐ Rectum ☐ Caecum ☐ Other (please specify)

Staging Type and Figures:
Stage:
T: N: M:

Dukes:
□ A □ B
□ C □ D

Site of Metastasis:

Grade:

☐ Grade I (Well Differentiated) ☐ Grade II (Moderately Differentiated)

☐ Grade III (Poorly Differentiated) ☐ Unknown

Morphology:

☐ Adenocarcinoma ☐ Mucinous (colloid) adenocarcinoma

☐ Signet-ring cell carcinoma ☐ Other (please specify)……...

Treatment:
1. Surgery: Date: _ _/ _ _/ _ _ _ _ (DD/MM/YYYY)

☐ Local excision or simple polypeptomy ☐ Resection and anastomosis

☐ Resection without anastomosis ☐ Other (please specify) ……

Tumour Size in cms: ☐ Number of nodes removed/ affected: ☐/☐

2. Chemotherapy: ☐ Yes ☐ No Date: _ _/ _ _/ _ _ _ _ (DD/MM/YYYY)

If, Yes a. ☐ FOLFIRI (5-fluorouracil, leucovorin and irinotecan)
b. ☐ FOLFOX (folinic acid (leucovorin), 5-FU, Oxaliplatin)
c. ☐ Other combination (please tick all drugs applicable)

☐ Folinic Acid ☐ Capecitabine ☐ Irinotecan

☐ Oxaliplatin ☐ Bevacizumab (Avastin) ☐ Cetuximab

☐ Panitumumab ☐ Unknown ☐ Other (please specifiy)……...

3. Radiotherapy: ☐ Yes ☐ No ☐ Not known

Supplementary Figure S4: Page Two of Colorectal cancer questionnaire (CRCQ) sent to all treating clinicians requesting confirmation and histology reports