DEVELOPING AND EVALUATING EXPERTISE IN COLONOSCOPY

Dr Kinesh Pradip Patel

BA(Hons) MB BS MRCP(Gastro)

Submitted for the degree of

Doctor of Medicine (Research)

Imperial College London

Division of Surgery and Cancer

2017

Abstract

The quality and safety of colonoscopy have become of paramount importance with the worldwide expansion in the utilisation of this procedure, especially with the introduction of colorectal cancer screening in many countries. It is well known that colonoscopic performance varies significantly between practitioners.

This thesis explores the effects of assessment in colonoscopy and polypectomy on performance and practice as well as trying to define the factors which differentiate the very best practitioners from the clinically competent.

Until 2011, there was no formalised way of certifying polypectomy competence. We looked at the effects of the introduction of mandatory polypectomy assessment on trainee endoscopists applying for certification of competency in the United Kingdom. This work showed that documentation of polypectomy competence significantly improved after these changes were introduced.

A global survey of polypectomy practice was undertaken to evaluate international guidance on polypectomy skills training and how trainers deliver teaching on polypectomy around the world, as well as trainees' experience of gaining polypectomy skills. Significant variability in endoscopists' experience of polypectomy training was found with few formal national guidelines published.

This led to an evaluation of expert endoscopists who underwent an accreditation process with some similarities before commencing Bowel Cancer Screening (BCS). We examined whether it was possible to predict future performance from a single assessment and found that criteria

used to assess whether candidates were competent could not predict the best performers from those who passed.

Several hundred expert BCS colonoscopists were then monitored over a three year period to determine changes in performance over time and whether long-term performance could be predicted. The best predictors of performance over time for all metrics were initial performance.

In order to identify key features of expertise in endoscopy, experienced colonoscopists were interviewed to distinguish characteristics of true endoscopic experts. These interviews revealed the importance of both technical and non-technical skills in defining expertise.

Preface

This thesis comprises original work conducted at the Wolfson unit for endoscopy, St Mark's Hospital, London and has not been submitted to any other university. All work from other sources has been appropriately referenced.

The research presented was carried out under the supervision of Dr Siwan Thomas-Gibson (Imperial College, London and St Mark's Hospital), Mr Omar Faiz (Imperial College, London and St Mark's Hospital), Professor Matt Rutter (Durham University and University Hospital of North Tees) and Professor George Hanna (Imperial College, London and St Mary's Hospital).

Acknowledgements

I am grateful to many people for their assistance and support in the preparation of this thesis.

First and foremost thanks must go to my principal supervisor Dr Siwan Thomas-Gibson, without whose support, forbearance and encouragement this thesis would not have been possible.

My co-supervisors, Mr Omar Faiz, Professor Matt Rutter and Professor George Hanna gave sound advice and guidance in their respective areas of expertise and for this I am also grateful.

I also thank Dr Anna Pinto, who provided indispensable advice with regard to qualitative study design, gave patient instruction on interviewing techniques and helped with the data analysis of the interviews. I am also appreciative of the help Dr Adam Haycock gave on study design.

Paul Dunckley and Pete Rogers were instrumental in providing access to the JAG dataset. I am also grateful to Paul Bassett, who advised on the statistical methods used in this thesis.

I must also acknowledge the Bowel Cancer Screening Programme Research Committee, who gave permission to access the datasets used in this thesis as well as approval to conduct the qualitative work interviewing screening colonoscopists.

Finally I would like to thank the participants in the research, who gave up their own time willingly and without hesitation.

Copyright Declaration

Table of contents

List of Tables

List of Figures

Abbreviations

AAFP	American Association of Family Physicians
ADR	Adenoma detection rate
ASGE	American Society for Gastrointestinal Endoscopy
BCS	Bowel Cancer Screening
BCSP	Bowel Cancer Screening Programme
BSG	British Society for Gastroenterology
CIR	Caecal intubation rate
COPD	Chronic obstructive pulmonary disease
DOPS	Direct Observation of Procedural Skills
DOPyS	Direct Observation of Polypectomy Skills
DNA	Deoxyribonucleic acid
EMR	Endoscopic mucosal resection
ESD	Endoscopic submucosal dissection
ESGE	European Society for Gastrointestinal Endoscopy
JAG	Joint Advisory Group for Gastrointestinal Endoscopy
KPI	Key performance indicators
MAP	Mean number of adenomas detected per patient
MCQ	Multiple choice question
MUPS	Munich Polypectomy Study
NCWT	Negative colonoscopy withdrawal time
NED	National endoscopic database
NPS	National Polyp Study
SSP	Specialist screening nurse practitioner
TNM	Tumour node metastasis
WEO	World Endoscopy Organisation

1 Introduction

1.1 Overview

The last decade has seen a true upsurge in the utilisation of colonoscopy, both as an investigative test to aid earlier diagnosis and as a modality to screen populations around the world for colorectal cancer. This rapid expansion in the provision of colonoscopy has led to a justifiable focus on the quality and safety of the procedure, especially as these have been convincingly linked with clinical outcomes[1,2]. It is known that considerable variability exists in performance at colonoscopy, some of which is undoubtedly due to technical ability. This thesis explores the effects of assessment in colonoscopy and polypectomy on performance as well as trying to define the factors which differentiate the very best practitioners from the clinically competent.

1.2 Colorectal cancer

1.2.1 Epidemiology

Colorectal cancer is common and a major cause of cancer related mortality in the developed world. In 2014, there were 41,265 cases of colorectal cancer recorded in the United Kingdom[3]. It is the second most common cause of cancer related mortality in the United Kingdom, with 15,903 deaths recorded in 2014. In 2010, the lifetime risk of developing bowel cancer was estimated by Cancer Research UK at 1 in 14 for men and 1 in 19 for women[3]. Survival was modest with about 57% of adult patients with colorectal cancer surviving for ten years or more[3].

These statistics are mirrored in other developed nations. In 2014, the National Cancer Institute in the United States of America estimated that there would be 136,830 new cases of colorectal cancer diagnosed, comprising 8.2% of all new cancer cases[4]. The Institute also predicted 50,310 deaths over this time within the USA[4].

It is thought that a significant proportion, of up to 40%, of the population in the Western world will develop adenomas at some point during their lifetime[5]. However, only a minority of these individuals will go on to develop colorectal cancer.

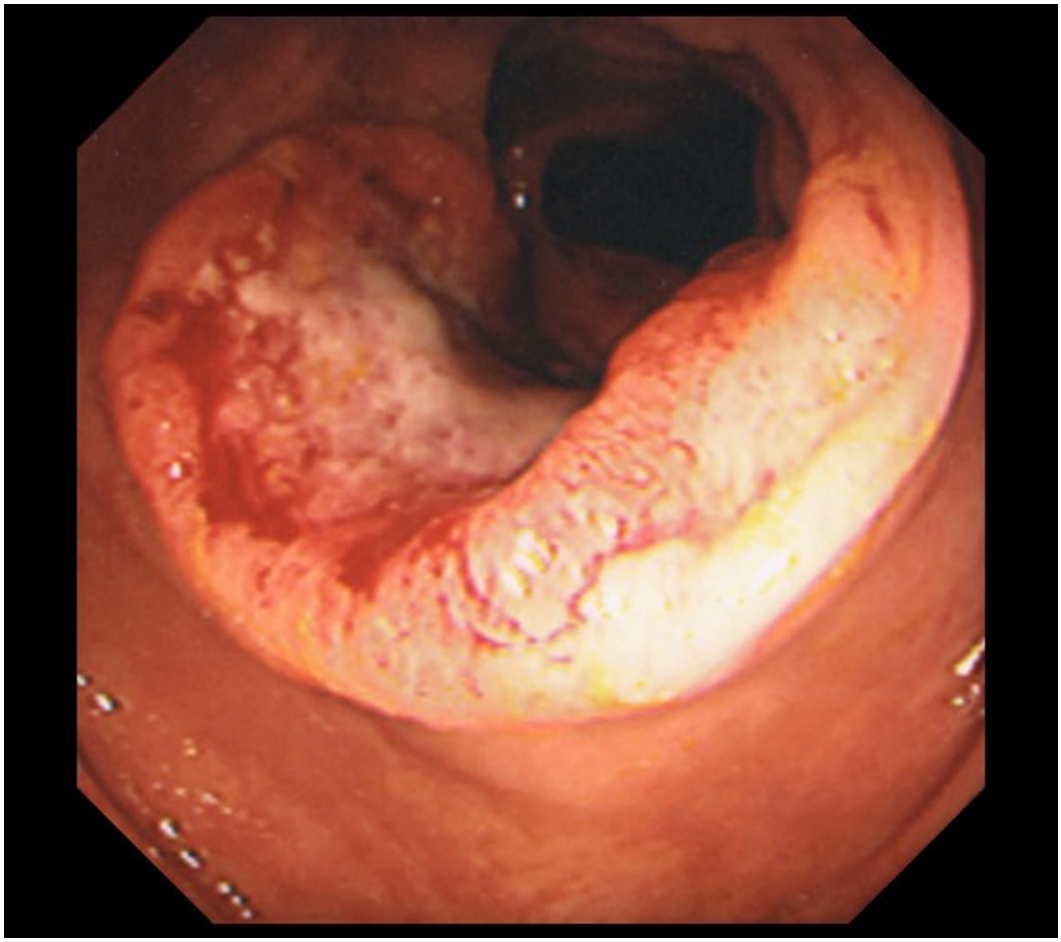

Figure 1.1. Endoscopic view of a colorectal carcinoma (image courtesy of Stephen Preston)

1.2.2 Colorectal cancer development

The majority of colorectal cancers develop over many years. In the 1960s, the now well-recognised stepwise progression from normal colonic epithelium to adenoma and then to carcinoma was first described[6] in the adenoma-carcinoma sequence.

Since then, there has been considerable effort devoted to trying to elucidate the specific genetic causes responsible for the development of both the hereditary colorectal carcinoma syndromes and those involved in sporadic colorectal cancer. A variety of mutations in DNA have been identified, including APC, KRAS, p53, TGF-β pathway mutations and features such as hyper-methylation implicated in the transition from adenoma to sporadic carcinoma[7]. This pathway is responsible for the majority of colorectal cancers.

However, in the 1990s interest grew in an alternative aetiology for predominantly right sided cancers. The serrated neoplastic pathway originates in similar precursors to hyperplastic polyps, although there are significant genetic differences, including high levels of microsatellite instability, CpG island methylation and BRAF mutations[8]. It is thought that this pathway may progress more quickly than the adenoma carcinoma sequence outlined above, although this remains contentious[9].

1.2.3 The role of screening in the interruption of natural history

Medical conditions worthy of consideration for screening were first described by the World Health Organisation in 1968 (Table 1.1).

Table 1.1. Wilson and Jungner criteria for screening[10]

Knowledge of disease
The condition should be important.
There must be a recognisable latent or early symptomatic stage.
The natural course of the condition, including development from latent to declared disease, should be adequately understood.
Knowledge of test
Suitable test or examination.
Test acceptable to population.
Case finding should be continuous (not just a 'once and for all' project).
Treatment for disease
Accepted treatment for patients with recognised disease.
Facilities for diagnosis and treatment available.
Agreed policy concerning whom to treat as patients.
Cost considerations
Costs of case finding (including diagnosis and treatment of patients diagnosed) economically balanced in relation to possible expenditures on medical care as a whole.

Colorectal cancer screening fulfils these criteria for several reasons. The long latent phase where polyps are present without overt neoplasia offers the opportunity to intervene and prevent the development of colorectal carcinoma at a later date. It is also clear that detection of cancer at an earlier stage is of considerable benefit to both the patient in terms of increased life expectancy (Table 1.2) and the health economy due to a lower cost of care.

Table 1.2. Colorectal cancer presentation by stage 1996-2002

Dukes' stage	TNM stage	Proportion of patients with colorectal cancer/%	5 year survival / %
A	T1/2N0M0	9	93
B	T3/4N0M0	24	77
C	Any T, N1/2, M0	24	48
D	Any T, any N, M1	9	6
Unknown		34	35

1.3 The role of polypectomy

1.3.1 Benefits

The attraction of polypectomy stems from the biological plausibility of interrupting the progression to cancer. The enthusiasm for removing polyps however has to be tempered by data suggesting that fewer than 5% of adenomas ultimately progress to cancer[11]. Nevertheless, consensus guidelines from several international organisations suggest that all adenomas are removed at screening colonoscopy as it is not possible to definitively predict which lesions are most likely to progress.

The National Polyp Study (NPS)[12], a retrospective cohort study in the USA published in 1993 compared 1418 patients who underwent polypectomy in the 1980s with three separate groups of patients: those who had had polyps detected on barium enema but declined surgical resection, a group from St Mark's Hospital who had had rectal polypectomy between 1957 and 1980 and a registry cohort monitoring people at average risk for colorectal cancer in United States. The incidence of cancer was at least 66% lower than that seen in the

comparator cohorts after six years and no patient died of colorectal cancer in the principal group.

This NPS group then underwent prolonged surveillance after the initial polypectomy to give a 37,073 person-year follow-up[13]. The median follow-up was 15.8 years with a longest follow-up of 23 years. Mortality was lower in the polypectomy group when compared to an age, sex and race matched cohort (standardised mortality ratio, 0.85; 95% confidence interval [CI], 0.81 to 0.90).

Twelve patients in the polypectomy cohort died from colorectal cancer compared to 25 predicted deaths in the matched population. The authors estimated the disease-specific mortality reduction from polypectomy to be approximately 53%. Mortality was comparable to an internal control group with a negative colonoscopy during which no polyps were found. These retrospective data suggested that the mortality reduction was durable for approximately 10 years after the initial polypectomy.

Other studies have tried to quantify the effect of colorectal cancer screening with colonoscopy. Only a minority of average-risk patients screened with colonoscopy will have colorectal cancer diagnosed on the index colonoscopy. The principal means therefore by which colonoscopy exerts its ability to reduce mortality is thought to be through polypectomy.

A case-control observational study from Germany compared 1688 patients with colorectal cancer and 1932 control participants. Colonoscopy in the 10 years prior to the study was

associated with a 77% risk reduction in the malignancy, with an effect across all cancer stages and age groups, except right-sided cancer in patients between 50 and 59.

1.3.2 Evidence of harms

Polypectomy itself is of course not a panacea. There are well documented risks associated with the technique, including bleeding, abdominal pain, colonic perforation, post-polypectomy syndrome and occasionally death[14,15].

A large prospective study in Munich examining complications and risk factors in 4000 snare polypectomies revealed an overall 10% complication rate, the majority of which were rated as minor[16]. However, 2.7% of patients experienced major complications[16].

Bleeding is the most common complication after polypectomy. Quoted rates in the literature vary significantly as there are many patient and polyp factors implicated in the risk of bleeding, as well as differences in the definition of bleeding[17-19]. The most important risk factors are likely to be polyp size and location, with a caecal or right colonic location associated with a significantly higher risk of bleeding[18,20]. Other risk factors associated with bleeding include polyp morphology, comorbidity (including diabetes, stroke, COPD and heart failure) and the use of oral anticoagulants or antiplatelet agents[21].

Data from the English Bowel Cancer Screening Programme (BCSP) examined 167,208 polypectomies in 130,831 patients. The overall bleeding rate in this study was 0.65%, with only a small minority of patients (0.04%) requiring transfusion[20].

The most feared sequela of colonic polypectomy is perforation. Perforation rates again vary depending on polyp size, morphology and colonic location with the caecum being a particularly high risk location for polypectomy. Data from the BCSP showed a 12 fold increase in the risk of perforation when sessile lesions were removed from the caecum, although the absolute number of events in the cohort was small due to a perforation rate of 0.06%[20].

These studies however included only patients undergoing screening colonoscopy. This patient group is known to vary from those presenting with symptoms warranting investigation in terms of having less comorbidity[21]. A community-based study from Northern California relying on data from 35,945 colonoscopies found an overall serious complication rate of 5%, with a perforation rate of 0.9%[19]. Surgery or transfusion for bleeding was required in 0.9% of patients. Polypectomy with either cold forceps or snare was associated with a ninefold increase in the risk of serious complications compared to purely diagnostic colonoscopy. Even in this study, the vast majority of participants (96%) were board-certified gastroenterologists performing over 150 colonoscopies per annum.

1.3.3 Effects of training

Those performing screening colonoscopy in centres submitting data for publication are more likely to be more experienced colonoscopists in addition to performing higher numbers of procedures annually. There is no direct literature on whether training endoscopists in polypectomy skills has an impact on complication rates.

However, a recent study has tried to address this question indirectly by comparing polypectomy techniques by endoscopists' specialty and experience and then correlating

these criteria with adverse events by using a dataset comprising 323,585 polypectomy procedures and 860 endoscopists across the entire state of Florida, USA[22].

Interestingly, there was considerable variation in polypectomy technique between different specialties and also between those performing high and low volumes of endoscopy.

Those performing low volumes of colonoscopy, defined as under 150 cases per annum, were more likely to use cold biopsy forceps to perform polypectomy as well as using hot biopsy forceps or ablation. Both of these techniques have become less favoured in experienced hands recently due to data suggesting that incomplete resection is more common with cold forceps[23] and thermal injury is more common with hot biopsy[24]. More advanced – and technically challenging - practices, such as snare polypectomy, were more commonly used by those performing over 300 cases per annum.

1.4 Variability in performance

1.4.1 Quality of colonoscopy

The introduction of screening programmes for colorectal cancer in many countries around the world over the past few years has brought the concept of the high-quality colonoscopy to the fore. The potential risks of introducing an invasive, uncomfortable and, to some, embarrassing test to an often otherwise healthy population in an effort to decrease mortality from colorectal cancer has meant that stringent monitoring of colonoscopists has been instituted in England to prevent harm to those being screened.

In order to assure the quality of any test, reliable outcome data are needed and screening colonoscopy offers the opportunity to collect such robust data prospectively in a relatively homogenous population. In the past, collection of such data in symptomatic patients was fraught with difficulty due to different indications, population types and electronic infrastructure necessary for large-scale monitoring of performance.

1.4.1.1 National

A prospective study looking at colonoscopy practice in the United Kingdom in 2004 revealed significant underperformance in the provision and practice of colonoscopy across the country[15]. Data were provided from 68 units on 9223 procedures, performed by 234 colonoscopists.

The adjusted caecal intubation rate (CIR) was 56.9% with a perforation rate of 0.13%. Only 19% of units recorded a mean caecal intubation rate of 90%. Just 17% of colonoscopists recalled having received training during their first 100 procedures. Discredited techniques such as transillumination and right sided abdominal finger indentation were widely used to identify the caecum. About one in 80 patients were admitted to hospital within 30 days of their colonoscopy.

A national strategy of training and quality improvement was agreed to address these shortcomings and colonoscopy practice was then re-audited in 2011. Data were collected on 20,085 colonoscopies performed by 2691 colonoscopists from 302 units. The unadjusted CIR was 92.3% and perforation rate 0.04% (against 56.9% and 0.13% respectively in the previous audit) demonstrating a remarkable improvement in practice in a comparatively short interval[2].

1.4.1.2 International

The findings in the United Kingdom were understandably a concern when initially reported. Published international comparators to this data also revealed considerable variability in performance.

A study aggregating data from seven teaching hospitals across the United States and Canada, including a military hospital was published a year before the first audit from United Kingdom[25]. Data were collected on 17,868 colonoscopies performed by 69 endoscopists. Just over half (55%) of endoscopists within the study cohort attained a CIR of greater than 90%. Nine per cent of endoscopists were recorded as having a CIR of less than 80%.

Similar findings were seen in an Italian study prospectively evaluating quality measures in four units. In total, 603 colonoscopies were performed. Of these procedures, 19.4% were incomplete, but in one centre the completion rate was only 56%[26].

1.4.2 Quality of polypectomy

1.4.2.1 Incomplete resection

The most serious long-term complication of incomplete polypectomy is the development of a colorectal cancer which could have been prevented. A proportion of interval cancers have been linked to previous incomplete polypectomy[27-29]. Few studies have however assessed factors affecting incomplete resection of polyps.

Patients presenting for outpatient colonoscopy at an academic medical centre in the USA were entered into a study if a sessile polyp between 5 and 20 mm was detected and hot

snared during colonoscopy[30]. All endoscopists were experienced in colonoscopy. After the endoscopist felt that the polyp had been completely removed, biopsies were taken from the margin of the polypectomy site and analysed by a specialist gastrointestinal pathologist for adenomatous tissue.

The incomplete resection rate varied significantly between endoscopists from 6.5% to 22.7%, with a mean of 10%. Larger polyps and those with serrated histology were less likely to be adequately resected.

However, it is not only larger polyps that can lead to incomplete resection. Cold forceps polypectomy is a common technique for the removal of smaller lesions in the colon. A prospective study assessing patients with diminutive polyps (less than 5 mm), assessed the effect of removing any visible polyp tissue with cold biopsy forceps[31]. The site of the polypectomy was then resected using endoscopic mucosal resection. Only 39% of diminutive polyps were resected completely using cold forceps. Again, the success rate for complete resection was significantly less for hyperplastic polyps compared to adenomas (odds ratio 5.1).

1.4.2.2 Complications

The ESGE published a guideline in 2012 on quality in screening colonoscopy [32], which referred to the minimum experience of screening colonoscopists. It suggested a link between endoscopist experience and polypectomy outcomes, based on a Canadian study which found that the risk of perforation and bleeding was increased by a factor of three with colonoscopists who performed fewer than 300 colonoscopies per annum [33].

This study assigned endoscopists into quintiles by the number of procedures performed during the one year study period. The lowest three quintiles had a significantly higher odds ratio (Figure 1.2) for post-colonoscopy bleeding or perforation than those in the highest quintile (greater than 379 procedures).

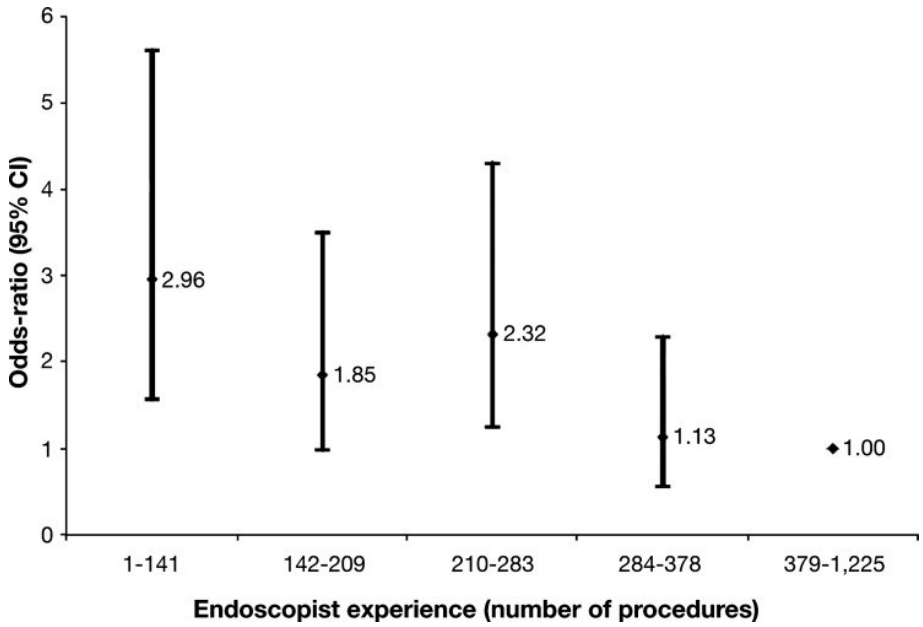

Figure 1.2. Odds ratio for bleeding or perforation by endoscopist colonoscopy volume[33]

The Munich Polypectomy Study examined risk factors for complications after 4000 colonic snare polypectomies. The principal risk factor was polyp size, with both minor and major complications increasing as larger lesions were tackled. A right-sided polyp location was also associated with poorer outcomes[16].

Analysis of adverse events within a much larger study in the BCSP has shown the findings with both bleeding and perforation strongly associated with polyp size. The caecum was also identified as the most risky part of the colon to perform polypectomy, although in this study

the overall complication rate was low with a bleeding rate requiring transfusion of 0.04% and perforation rate of 0.06%[20].

The inference from these studies is that inexperienced endoscopists tackling larger, particularly right-sided, lesions are likely to have significantly higher complication rates.

1.4.3 Implications of variability in performance

1.4.3.1 Incomplete examination

Safe examination of the entire colon is the first hallmark of a high quality examination. It does not seem biologically plausible for colonoscopy to reduce right-sided malignancy without an inspection of the right colon combined with removal of precursor lesions. However, a recent randomised trial of flexible sigmoidoscopy screening has shown a 10 per cent reduction in the rate of proximal colorectal cancer[34]. This may be due to identification of high-risk lesions in the left colon which then lead to colonoscopy and the detection of advanced proximal polyps.

A case-control study from Ontario, Canada evaluated the association between colonoscopy and death from colorectal cancer by matching 10,292 patients with 51,460 controls. There was an inverse association between colonoscopy and death from left-sided colorectal cancer but no survival benefit was seen in the right side of the colon[35].

1.4.3.2 Missed polyps

The implications of missed polyps are twofold. Firstly the detection and subsequent removal of polyps is necessary to interrupt the progression outlined above from benign polyp to carcinoma. Secondly, it is widely accepted that the smaller lesions detected by endoscopists

with higher adenoma detection rates are less likely to progress to overt malignancy. However, the detection and removal of polyps is likely to be associated with a more careful examination and a lower risk of interval cancer, as outlined below.

Even in experienced hands, missing polyps is not unusual. A study evaluating the time spent by endoscopists who had performed over 3000 lifetime colonoscopies inspecting the colonic mucosa on withdrawal found significantly different ADRs between endoscopists. A withdrawal time of greater than six minutes was associated with a significant increase in the number of adenomas detected but the relationship was largely linear with no threshold time seen above which patients derived no benefit in terms of increased adenoma detection[36].

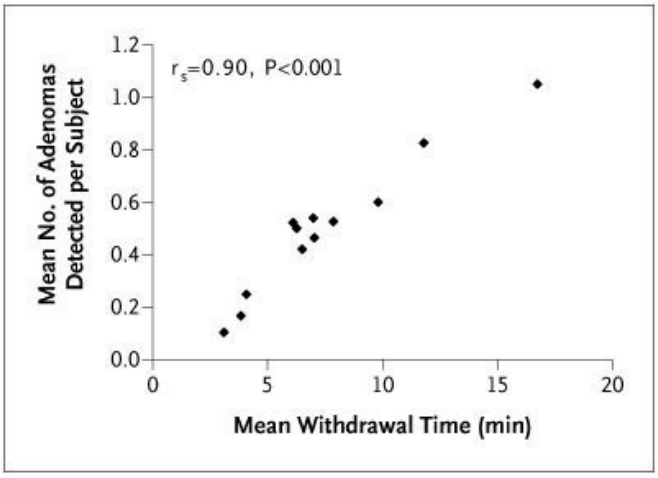

Figure 1.3. Withdrawal time against mean number of adenomas detected per patient[36]

Similar findings have been seen in the BCSP[37]. However, other large studies have shown withdrawal time not to be strongly associated with polyp detection[38]. The most likely explanation for this is the importance of the quality of inspection. Given there is no validated

method of assessing this, the time taken to inspect the colon has been widely adopted to act as an imperfect surrogate marker for inspection quality.

1.4.3.3 Post colonoscopy cancer / interval cancer

There is considerable debate in the literature about the precise definition of post-colonoscopy cancer. It is thought that the majority of colorectal cancer presenting within a few years after colonoscopy is due to missed lesions rather than the rapid evolution of a *de novo* carcinoma.

One study, which obtained data from a large Canadian registry, found that 7.9% of all patients with colorectal cancer had a colonoscopy between 6 and 36 months before their diagnosis of malignancy[39]. Areas such as the splenic flexure, which are notoriously difficult to visualise completely, were associated with higher rates of missed malignancy. There was also a significant association between higher rates of post-colonoscopy cancer and a non-gastroenterologist endoscopist.

A seminal paper analysed 45,026 subjects from Poland at average risk of colorectal cancer undergoing screening colonoscopy[1]. In this study, any diagnosis of colorectal cancer between the time of screening colonoscopy and the next scheduled surveillance colonoscopy was included as an interval cancer. The median follow-up time was 52 months. There were a total of 42 interval cancers, the majority (92.9%) of which occurred in patients who had reportedly had a normal colonoscopy with no adenomas detected.

An adenoma detection rate (ADR) within the programme was calculated for each of the 186 endoscopists. The endoscopists were then categorised into separate groups, with the highest performing group having an ADR of greater than 20%. There was a significant association

between endoscopists with an ADR less than 20% and an increased hazard ratio for the development of interval cancer.

More recently, a similar study identified 264,972 colonoscopies in Northern California, USA performed in 252,842 patients[36]. There were 712 interval cancers, as defined by cancer diagnosed between six months and ten years after the colonoscopy. The median delay between the index colonoscopy and the diagnosis of colorectal cancer was 39 months. Two thirds of the interval cancers were detected more than three years after the initial procedure.

Again, the adenoma detection rate for each individual screener was calculated and endoscopists placed in quintiles (Table 1.3), which were then correlated against the risk of interval cancer.

Table 1.3. Adenoma detection rate in each quintile[36]

Quintile	ADR range / %
1	7.35-19.05
2	19.06 – 23.85
3	23.86 -28.40
4	28.41 – 33.50
5	33.51 – 52.51

This study showed a continuously diminishing risk of interval cancer with increasing adenoma detection rates (Figure 3). Again, there was considerable variability in adenoma detection rate with a sevenfold difference seen between the worst and best performers.

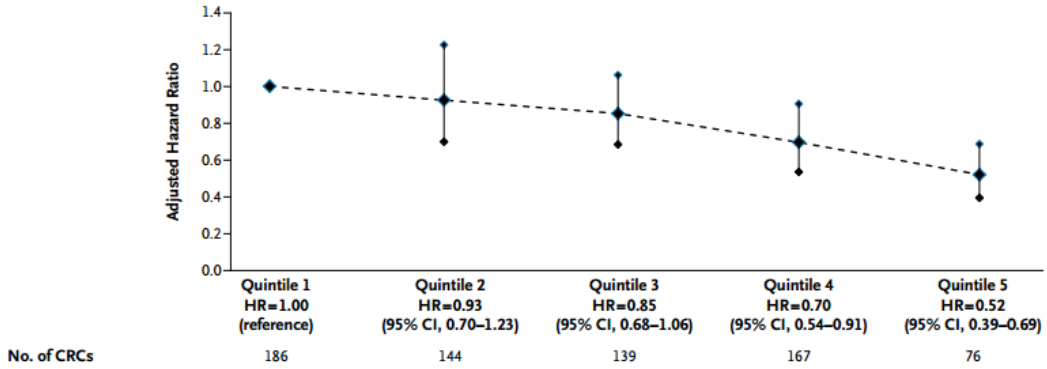

Figure 1.4. Hazard ratio for colorectal cancer by quintile of colonoscopist adenoma detection rate[36]

A further study encompassing the entire population of Utah, USA sought to examine characteristics of interval cancers in the state[40]. A total of 159 interval cancers were identified from 126,851 patients. The majority (55%) of interval cancers were in the proximal colon, with 20% at the rectosigmoid junction. Given left-sided colorectal cancer is more common in the general population, right sided interval cancer seems to be overrepresented. This may be due to well-described technical difficulties visualising the right side of the colon[41] in those less skilled at colonoscopy, different tumour biology or a combination of both of these factors.

1.5 Training to expertise

1.5.1 Training in colonoscopy

All the major organisations representing endoscopy have recognised the value of training, not only to ensure that current standards are maintained but to iteratively improve each new cohort of trainees to surpass the last.

There has been much debate in the literature about the most effective ways to train endoscopists to perform colonoscopy. Training itself is recognised to be heterogeneous in its nature[42], encompassing a gamut of experiences, from high quality trainee trainer interactions to passive learning.

1.5.1.1 Completion of procedure

The best studied aspect of colonoscopy, with respect to training, is procedural completion, that is insertion of a colonoscope to the caecum. This has been used as a surrogate marker of competency for many years as it is easy to measure objectively and is not influenced significantly by patient type. An unadjusted caecal intubation rate of 90% is generally accepted as a level signifying competence. This metric permits no more than one in ten incomplete procedures due to obstructing lesions and poor bowel preparation. The unadjusted CIR has gained favour over the adjusted CIR as it does not permit gaming of the statistic by the endoscopist by, for example, over-attributing poor bowel preparation as the principal reason for failure.

Most studies to date have centred on trying to calculate the number of colonoscopies trainees need to undertake to reach a caecal intubation rate of 90%.

Spier et al [43] performed a retrospective analysis on trainees performing colonoscopy at a single centre. The study included 11 trainees, performing 770 procedures over 18 months. The specific aim was to examine how many procedures were necessary for trainees to reliably attain a caecal intubation rate of 90%. All trainees managed this after 500 procedures but no trainee attained this level after the guideline suggested 140 procedures.

A further paper by the same author [44] looked at surgical residents with very little endoscopy experience participating in a two-month intensive endoscopy rotation. Trainees performed on average 80 ± 35 colonoscopies (range, 40–160) during the rotation. 71% did not attain a CIR over 50%. No trainee managed to attain a CIR of 90%.

A study from the Cleveland Clinic examined CIR and time to caecum for 18 trainees over five years[45]. The mean CIR for the first 25 cases was 43.1%, rising to 75.1% for cases 100-125.

Selvasekar et al sought to look at performance data for six colorectal surgery fellows in the United Kingdom to try to ascertain the volume of procedures needed to attain a CIR of 80%. The data showed that 114 procedures were needed to reach this level of competency although the number needed to reach 90% was not recorded.

A study performed at the Mayo Clinic [46] looked at skills acquisition over time by trainees, by means of a novel colonoscopy skills assessment tool devised at the hospital. The subjects were 41 trainees performing 6390 procedures over three years, with 3936 assessment forms completed. On average, 275 procedures were needed to attain competence using this assessment tool. After performing this many procedures, the CIR was on average 85%.

A recently published Korean study examined factors affecting caecal intubation failure in trainees[47]. The subjects were four novice trainees. In the first 50 cases, the CIR was 62%, growing to 93% after 250 cases had been successfully completed.

Lee at al performed a prospective multicentre trial in 15 tertiary academic hospitals, looking at the performance of 24 first year GI trainees performing 4351 colonoscopies [48]. The prespecified outcomes were CIR over 90% and a caecal intubation time of less than 20

minutes. 150 procedures were needed to attain a CIR of 91.3% but there were continuing improvements documented up to 300 colonoscopies, when the caecal intubation rate reached 98.7%. The mean time taken to reach the caecum also fell with experience.

A large prospective study was published in 2001, examining 13,580 procedures performed between 1998 and 1999 [49]. The endoscopists were principally highly experienced surgical staff with a median endoscopy experience of over 20 years and over 1000 procedures. Outcomes were self-reported. 10% of endoscopists were trainees with fewer than 50 procedures documented; 6.5% of procedures were performed by those who had carried out 100 to 200 colonoscopies. Those who performed fewer than 50 procedures had a CIR of between 75.9% and 84.4%. This rose to 91% for those who performed 50 to 99 procedures. In this study, a minimum procedure volume of 100 colonoscopies per year was necessary to attain a CIR greater than 90%.

The most recent and highest quality data on the learning curve for colonoscopy come from the electronic database of all trainees in the United Kingdom. This is the largest study in the literature addressing this specific question. Analysis of 36,730 colonoscopies performed by 297 trainees showed that this cohort reached a CIR of 90% after a mean of 233 procedures. Only 41% of trainees attained the same standard after 200 procedures[50].

These studies all have similar findings: a variable number of procedures is needed to attain competency in diagnostic colonoscopy, but this figure is likely to be in excess of previous recommendations. As a result of these data, many national societies have updated their own guidance on training and several, including the United Kingdom, have moved away from numbers-based training to competency based assessment (section 1.5.1.2.2).

It is likely that the variation in the data pertaining to the learning curve for diagnostic colonoscopy is multifactorial. Individual variation in skill acquisition is inevitable whenever humans try to acquire a new technical ability[51]. Other external factors likely to impact upon the rate of skill acquisition include the quality of training, including mentorship and feedback, as well as the intensity of cases.

Despite the many studies described above trying to determine the exact number of cases needed to achieve competency, in reality the identification of a predetermined number is more useful in terms of highlighting when a trainee is likely to be fit for a summative competency based assessment. There is clear evidence from the literature that numbers alone do not confer competency on endoscopists, with marked variation in performance metrics in even those that have completed several thousand procedures[36].

1.5.1.2 Polypectomy

1.5.1.2.1 Technique

Polypectomy is often described as a single defined task. However, there is still considerable debate in the literature about the best way to remove polyps from the colon. Previous work has tried to gain a consensus on polypectomy technique with some success but significant differences between experts remained in some areas[52].

The principal factors affecting decision-making about how best to remove polyps include size, location in the colon and morphology. Different techniques are applicable for sessile or flat polyps compared to pedunculated lesions. The smallest lesions, those up to 3mm may be removed with cold forceps polypectomy. This technique used to be much more prevalent

even for larger polyps but the recent data discussed above suggest that even with diminutive lesions, there is a significant risk of incomplete resection. The advantages of polypectomy with cold forceps include no risk of diathermy injury and a negligible risk of perforation[53].

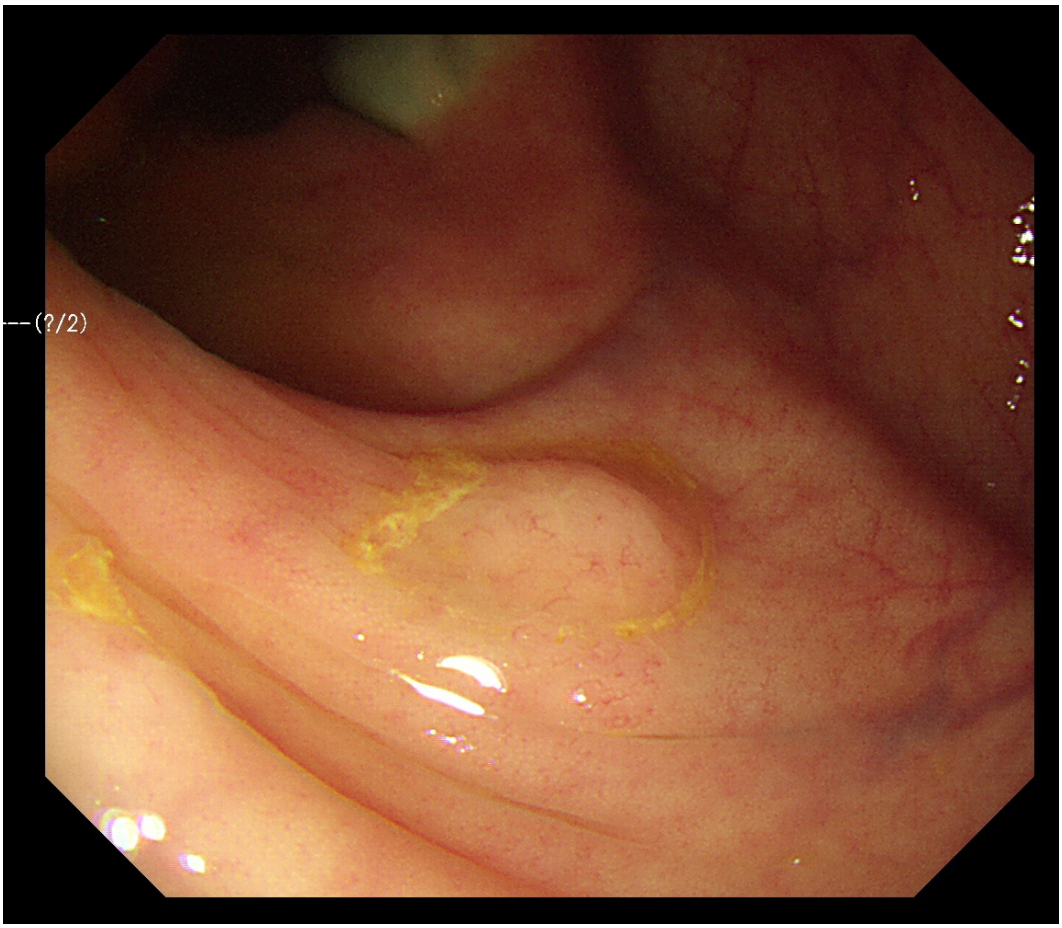

Figure 1.5. A diminutive colonic polyp, typically resected with cold forceps (image courtesy of Stephen Preston)

In the past, hot forceps polypectomy was widely used to try to address the problem of incomplete resection. The theory is attractive in that any residual tissue at the base of the polyp is destroyed with electrocautery, while still providing a histological specimen for analysis. Concerns about safety and incomplete resection[54] have led to a decrease in utilisation of this technique over time[55,56].

The most commonly used method of removing colonic polyps is snare polypectomy, the utilisation of a metal loop that is first opened in the colonic lumen over the polyp and then closed once the polyp and a rim of normal tissue has been encapsulated within the snare. The polyp can be resected either with or without the application of a diathermy current through the snare.

Snare polypectomy itself is technically demanding and is a complex multistep procedure, including assessing the lesion to be resected, optimising positioning of the endoscope, accurately directing the snare around the polyp ensuring the correct amount of tissue is trapped within it and applying the correct duration of diathermy whilst closing the snare. Once the polyp is resected, polyp retrieval, ensuring that no visible residual tissue remains and dealing with complications, such as bleeding or colonic perforation, are all critical aspects in ensuring a safe and effective procedure. This procedure, with its numerous steps, has been effectively summarised into a competency framework, the DOPyS (section 1.5.1.3.1.4), to permit assessment of this technique.

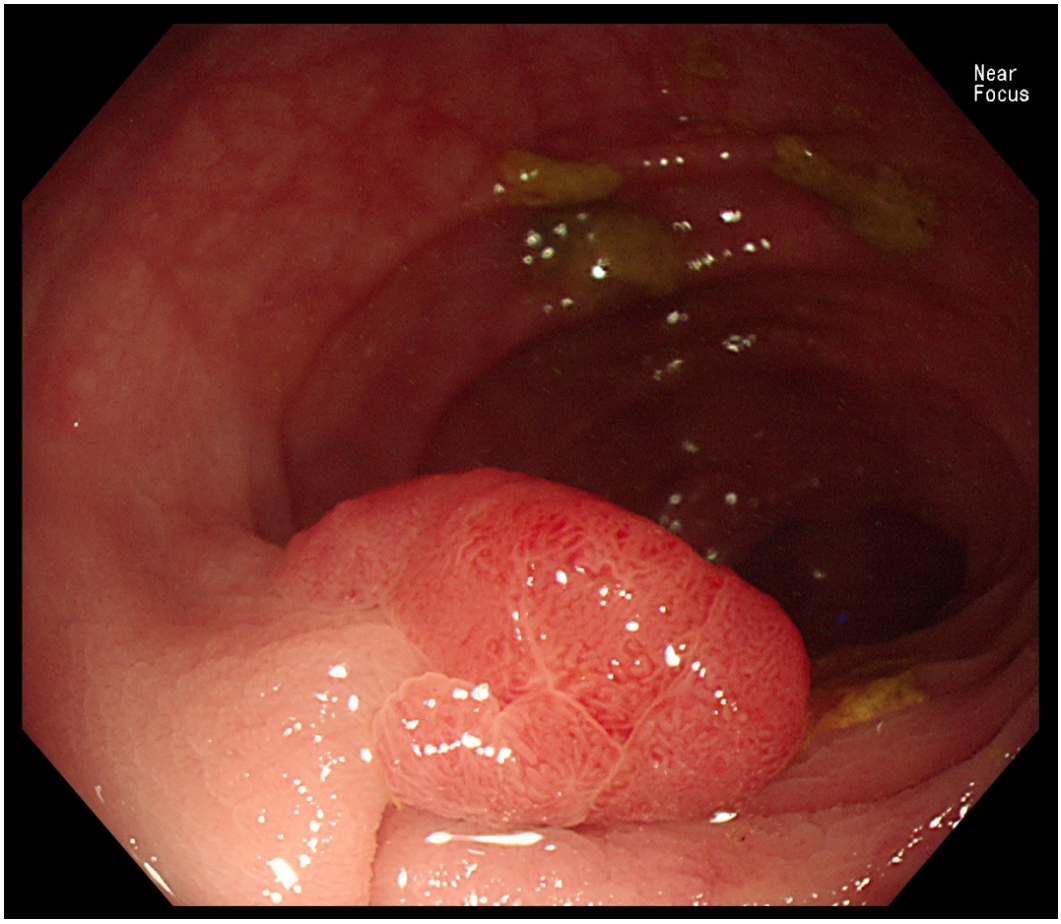

Figure 1.6. A pedunculated colonic polyp, typically resected by snare polypectomy (image courtesy of Stephen Preston)

Endoscopic mucosal resection is utilised for those polyps that cannot be safely removed with a snare due to their morphology. This involves the injection of fluid into the submucosa to lift the polyp on a cushion of injectate to facilitate resection with a snare *en bloc* or in several pieces. The cushion helps prevent iatrogenic complications due to thermal burns[57]. Endoscopic mucosal resection can in this way be used to tackle larger lesions safely, but again is a demanding skill both with regard to technical ability and decision-making. These skills have traditionally been poorly taught.

There is a need to define the learning curve for polypectomy so that the acquisition of these skills is evidence-based and mirrors the broad literature which already exists for diagnostic colonoscopy completion. This work is already underway for endoscopic submucosal dissection[58-61] yet remains lacking for standard polypectomy.

1.5.1.2.2 Training

A competent polypectomy has two ultimate aspects: the polyp is completely removed and the procedure is undertaken safely without complications.

Whereas there are now many sources of data on training for colonoscopy completion, this is mirrored by a significant gap in the literature on polypectomy training. A study has validated a polypectomy skills simulator in an ex vivo porcine model and shown good correlation between DOPyS scores attained in the simulator and in patients[62] but there are no prospective studies examining the number or type of polypectomies that trainees need to perform on average to become competent.

However, endoscopic societies have in some cases outlined statements based on consensus opinion as to the skills that trainees should possess before being deemed competent at colonoscopy.

1.5.1.2.3 American Society for Gastrointestinal Endoscopy (ASGE)

The ASGE core curriculum, last revised in 2012[63], lists snare polypectomy as a core motor skill required for a trainee to be competent in colonoscopy overall.

The skills required to perform safe polypectomy are listed, including snare utilisation, diathermy settings and non-technical skills. Skills such as injection of the stalk, loop placement and clip placement for large pedunculated polyps are also recommended.

The guidelines state the trainees should have decision-making skills to allow them to identify polyps that cannot easily be resected with a colonoscope, such as those spanning more than one fold or spanning greater than one third of the luminal circumference of the colon. It is suggested that trainees avoid pedunculated polyps over 1.5 cm due to the risk of bleeding.

Further recommendations are made with regard to the post-polypectomy skills, including the management of complications, such as post-polypectomy bleeding and polyp retrieval techniques.

There are no numbers suggested before attaining competence in polypectomy over and above those recommended for general trainees in the latest revision of the ASGE guidelines (275 procedures).

1.5.1.2.4 American Association of Family Physicians (AAFP)

This guideline, published initially in 2000[64], was aimed at physicians practising in the community. It refers to older ASGE guidelines[65] recommending 100 diagnostic colonoscopies and 20 snare polypectomy procedures before competence is reached and highlights that this was based on expert opinion rather than any scientific measurement. However, the AAFP recommended that 50 cases as the first endoscopist is a reasonable level to attain basic competency.

An updated version of the guideline, published in 2013, retained the same volume guidance of 50 diagnostic colonoscopy procedures to attain competency, based on a review of several studies[66,67].

It defines competence as the acquisition of psychomotor skills for technical proficiency in performing the procedure in conjunction with the cognitive skills necessary to know when to perform the procedure and how to interpret findings and states that 'there is no scientific data correlating the volume of colonoscopies performed with the acquisition of competence'.

1.5.1.2.5 European Society for Gastrointestinal Endoscopy (ESGE)

The 2012 ESGE guideline has no specific recommendations on attaining competence in polypectomy[33]. However, it states that the annual number of polypectomies is likely to be more important than the annual number of diagnostic procedures in ensuring competence in polypectomy. It also quotes the German quality assurance programme's recommendation of an annual minimum of 10 snare polypectomies to maintain accreditation.

1.5.1.2.6 UK / JAG guidelines

The United Kingdom has been at the forefront of developing colonoscopy and polypectomy competency-based training and assessment. There are stringent guidelines to certify competence of those performing both diagnostic and screening colonoscopy. The processes involved are described in detail in section 1.5.1.3.1 below.

1.5.1.3 Role of assessment

A paper from the early 1980s describing the distortion of medical students' behaviours by a known change in the weighting of finals examinations first introduced the concept of assessment driving learning[68]. A new examination mark scheme aimed to emphasise the importance of ward-based learning and so students were assessed every four weeks on their performance. However, these clinical assessments were invariably rated satisfactory. The corollary of this was that students increased the amount of time spent studying in the library in an effort to improve their marks in the theoretical part of the examination, a perverse but logical reaction to the new assessment criteria.

1.5.1.3.1 Accreditation processes

In the United Kingdom, there are two separate levels of colonoscopy competency certification. These processes were established in response to the poor performance described in the 2004 audit of colonoscopy practice in United Kingdom.

1.5.1.3.1.1 Trainees

The first route applies to all trainee endoscopists, who wish to gain certification in colonoscopy. This process is mandated by the Joint Advisory Group for Gastrointestinal Endoscopy (JAG).

The requirements of certification have changed since its first introduction. Initially, trainees were required to complete 200 colonoscopies, providing some evidence of satisfactory formative assessment during that time.

Attendance at a basic skills in colonoscopy course was also mandatory. After these requirements have been met, trainees were asked to submit a copy of their written logbook with data pertaining to caecal intubation rate and average sedation rates, along with a summative assessment. This assessment comprised two witnessed procedures by two endoscopy trainers, using the Direct Observation of Procedural Skills (DOPS) tool described below. The portfolio and assessments were then verified centrally and a certificate of competency issued.

In 2011, these requirements changed reflecting the impact of technology on record-keeping and the perceived lack of polypectomy assessment in trainees' portfolios.

Trainees applying for certification since this time have been required to complete an electronic portfolio contemporaneously documenting each procedure including data on indication, completion, findings and interventions.

When a trainee attains certain prespecified key performance indicators, including an unassisted caecal intubation rate over 90% and satisfactory formative assessments of colonoscopy and polypectomy technique, they are eligible to apply for a summative assessment. Just as before, the trainee is watched performing two colonoscopies by two assessors and their performance assessed using the DOPS tool. Adequate performance at this stage permits provisional certification of colonoscopy.

At a later date, after further independent practice and when polypectomy assessments have been performed on lesions between 1cm and 2cm, an application can be made for full certification.

1.5.1.3.1.2 Bowel Cancer Screening Programme (BCSP)

The English BCSP has a unique, rigorous, multi-step accreditation process before colonoscopists are permitted to participate in the screening programme. All colonoscopists have to have performed over 1000 colonoscopies in their careers. Both an unadjusted CIR of over 90 per cent and polyp detection rate of over 20 per cent are mandatory in the procedures performed in the year before applying for accreditation. Independently verified data are also collected on complications within the previous year of endoscopic activity, including vasovagal attacks, bleeding, unplanned admissions after colonoscopy and the use of sedative reversal agents.

The individual then takes a 1-hour, single best answer, multiple choice question test of 30 questions, each with 5 stems, to demonstrate factual knowledge about colonoscopy and the programme, and then finally undergoes a formal structured assessment process during which they are observed performing two colonoscopies on patients by two specially trained assessors, who themselves are screening colonoscopists within the BCSP.

Performance during the two observed procedures is independently scored by both assessors using the DOPS tool. If any polyps are encountered, then colonoscopists are expected to undertake their usual clinical practice in removing the polyps if clinically appropriate. If polypectomy is performed, this too is assessed using the Direct Observation of Polypectomy Skills (DOPyS) tool described below.

1.5.1.3.1.3 DOPS

The DOPS tool was introduced into clinical practice in parallel with developments in medical education highlighting the importance of workplace-based assessments to certify competency in a wide variety of procedures[69]. Similar assessments have been introduced for all junior doctors and medical students, tackling skills as diverse as cannulation to cardiac catheterisation.

The colonoscopy DOPS was developed by a multidisciplinary panel with a particular interest in training in endoscopy[70]. The 20 domains assessed using the DOPS (Appendix 1) were developed iteratively, along with the descriptors for each domain which describe the level of attainment candidates need to demonstrate in order to score particular marks.

Each domain is scored between 1 to 4 (Table 1.4), with 1 and 2 being regarded as signifying further improvement is necessary and 3 and 4 denoting competent performance.

Table 1.4. DOPS levels of achievement for each domain

Level	Description
1	Accepted standards not yet met, frequent errors uncorrected
2	Some standards not yet met, aspects to be improved, some errors uncorrected
3	Competent and safe throughout procedure, no uncorrected errors
4	Highly skilled performance
N/A	Not applicable: item could not be assessed

The DOPS tool has been validated with generalisability theory in clinical practice within the Bowel Cancer Screening Programme, with the scores correlating well with expert

assessment[70]. Both candidates and examiners agreed that the DOPS was a valid way of assessing their ability.

Interestingly, a similar methodology using generalisability theory was used to validate DOPS to assess intubation skills in anaesthetic trainees. The results mirror those found in the colonoscopy validation with high levels of validity and reliability[71].

1.5.1.3.1.4 DOPyS

The DOPyS tool (Appendix 2) was devised in response to concerns that the standard DOPS assessment did not adequately certify competency in polypectomy technique[52]. The different skills necessary for polypectomy were identified by task deconstruction and a list of 34 parameters thought to be fundamental to the safe practice of polypectomy derived from this.

BCS colonoscopists were asked to record polypectomy videos and submit them for independent evaluation by seven experts. Videos were utilised rather than live assessment to both minimise the observer or Hawthorne effect, whereby behaviour changes due to knowledge of the observation, and to minimise logistical difficulties in assessing study participants on different days at different sites.

Consensus on the skills and competency level required to attain a pass mark for each of the previously identified parameters was sought. The same marking scheme used in the DOPS tool (Table 1.4) was also used for the DOPyS for clarity.

The videos were then scored independently by 7 blinded expert assessors and the results analysed to evaluate the level of concordance between them. By using generalisability theory, the validity of the tool was proven.

The desire to certify the competency of those undertaking bowel cancer screening underpinned the development of the DOPyS. Once the tool had been introduced, there was a legitimate call for trainees to also have to demonstrate competence in polypectomy.

The DOPyS was therefore instituted as an obligatory part of trainee certification and the electronic portfolio. However, it has never been validated in this cohort or indeed during the assessment of live polypectomy as its initial purpose was to evaluate whether the competence of endoscopists could be remotely assessed by using videos of procedures.

Those using the DOPyS tool within the Bowel Cancer Screening Programme have been trained to its use and the requisite standards necessary to attain a passing mark for each domain. During the initial work where the tool was devised, there was significant discordance amongst many parameters between expert endoscopists. However similar training has not been deployed to the far larger number of endoscopy trainers now tasked with assessing polypectomy nationally for hundreds of trainees. There is a concern, backed by the initial pilot data, that standards vary between trainers and because of this the assessment may lack external validity without dedicated training of assessors.

The development of the DOPyS tool has stimulated interest in polypectomy skills training. A recent study described the development of an ex vivo porcine simulator for polypectomy

training with the DOPyS used to assess the skill of endoscopists of differing experiences at performing polypectomy[62].

The same simulator was then used to allow endoscopists to rate their own polypectomy skills using the DOPyS. There was poor correlation between self-rated scores and those of independent experts with novices tending to score themselves too harshly and experienced endoscopists overestimating their own abilities[72].

With its widespread deployment, there is a need to understand how the DOPyS tool is being used in everyday practice and the effects that it has had on training.

1.5.2 Expertise

The concept of expertise has been widely discussed across many arenas of human endeavour. For many years, psychologists and other professionals have tried to elucidate the characteristics of experts in myriad fields, including medicine. Table 1.5 shows psychological characteristics of experts from two wide-ranging reviews of the field[73 74].

There is also considerable debate about the true performance of experts themselves. Some data suggest that experts are overly prone to errors due to utilisation of familiarity heuristics, their reliance on their familiarity with previous events when deciding on the most appropriate course of action[75]. However, other literature supports the concept of experts as a disparate class of individuals with different skills and thought processes to others[76].

Table 1.5. Psychological characteristics and strategies of experts

Characteristics **Shanteau, 1992[77]**	**Characteristics** **Glaser and Chi, 1988[78]**
Extensive and up to date content knowledge	Experts excel mainly in their own domain
Highly developed perceptual/Attentional abilities	Experts perceive large meaningful patterns in their domain
Sense of what is relevant when making decisions	Experts are fast (faster than novices at performing the skills of their domain) and they quickly solve problems with little error
Ability to simplify complex problems	Experts have superior short term and long term memory
Ability to communicate	
Handle adversity better	
Experts are better at identifying and adapting to exceptions	Experts see and represent a problem in their own domain at a deeper (more principled) level than novices; novices tend to represent a problem at a superficial level.
Self confidence in decision making	
Adapt decision strategies to changing task conditions	
Strong sense of responsibility and willingness to stand behind their recommendations	Experts spend a great deal of time analysing a problem qualitatively
Strategies (Shanteau, 1992)	
Willingness to make continuous adjustments in initial decisions	
Experts get help from others to make better decisions	Experts have strong self-monitoring skills
Experts often make use of formal or informal decision aids	
Experts make small errors and try to avoid making large mistakes	
They operate as though coming close is generally good enough	
Experts follow some sort of divide and conquer strategy	
Break problems down	

One study examined expertise in anaesthetists in a teaching hospital in North America. Eight cases of patient anaesthesia were recorded and 30 hours of conversations occurring during the cases were transcribed. These episodes were analysed and the physical actions and thoughts were classified and interlinked. Expert anaesthetists anticipated future problems from 'warning flags' in the history before the case had commenced and prepared for uncommon eventualities that might conceivably occur[79].

In specialties involving practical procedures, it is self-evident that technical skill is one of the hallmarks of expertise. However, the absolute importance of this ability has recently been questioned. One study showed that poor communication was responsible for 43% of errors made in surgery[80]. Cognitive and social skills, grouped together under the broad category of non-technical skills seem to play a critical part in the delivery of safe and effective care[81].

1.5.2.1 Colonoscopic expertise

Although broad concepts in expertise have been defined, there has been little research on experts in colonoscopy. Much research has been devoted to the attainment of high quality performance metrics such as CIR and ADR, yet the presence or absence of other psychological skills highlighted in Table 1.6 has never been studied in endoscopic experts themselves.

The most relevant study was published in 2014, after a Delphi survey was conducted amongst 8 expert endoscopists from the Netherlands[82]. Participants were asked to list factors that they deemed essential for high quality colonoscopy. The factors were then validated with the help of external panel. The final results from the Delphi survey are shown in Table 1.6.

Table 1.6. Final Delphi results

Efficacy/endpoints	**Safety/side effects/behaviour**
Knowledge	
Adequate identification of endoscopic image	Knowledge of own boundaries
Basic colonoscopy technique	Knowledge of material and options for polypectomy
Knowledge of complications and registration	Understanding and solving loops
Skills	
Caecal intubation rate	Skills and hand-eye coordination
Polyp detection and removal	Patience and precision
Competency in intervention techniques	Small, gentle movements
Assessment of mucosa	Minimising insufflation
	Proper position for intervention
	Feeling of equipment

This work has begun the process of trying to define endoscopist factors that lead to a high quality colonoscopy but has not been able to differentiate true expertise from clinical competency.

Previous work has tried to pinpoint some of the non-technical skills required to perform endoscopy to a high level by utilising semi-structured interviews[83]. Several overall domains were identified in four principal categories: communication and teamwork, situation awareness, leadership and judgement and decision-making.

These characteristics have significant differences to those postulated as being important in endoscopic expertise in the Delphi survey. It is likely that all these attributes are important to some extent but how those who perform at a significantly higher level than other endoscopists rank in each of these categories is unclear.

There is a need to define true expertise in colonoscopy given its now extensive use in a healthy population in order to try to close the gap between those performing at the very top of their field and those working at a lower, albeit entirely competent, level. As it is likely that there is no threshold effect for the benefits those undergoing bowel cancer screening are able to derive from endoscopists with higher adenoma detection rates, closing the performance gap would serve both the individual and the wider population through a decrease in the burden of cancer in the community.

1.6 Research questions

- How does the formal assessment of polypectomy impact training?
- How are endoscopists trained in polypectomy around the world?
- How does a structured accreditation affect the performance of endoscopists?
- How does colonoscopic performance vary between different endoscopists?
- How does colonoscopic performance vary over time?
- What constitutes colonoscopic expertise?

Chapter 2 describes the impact of the introduction of formalised polypectomy assessment on trainees and training in the United Kingdom and asks what have been the consequences of the introduction of DOPyS into trainee certification and whether this has helped determine the learning curve for polypectomy.

Chapter 3 includes a review of polypectomy training and assessment guidelines from around the world as well as an international survey of colonoscopy trainers and trainees. The research

asks about the state of polypectomy training worldwide and whether there is a consensus on how endoscopists should be taught this technique.

Chapters 4 and 5 examine the performance of screening colonoscopists in the English Bowel Cancer Screening Programme.

Chapter 4 focuses on the relation between performance during the initial accreditation and quality data from the first 12 months of screening activity for each endoscopist within the BCSP, asking whether it is possible to predict expert performance from an accreditation process.

Chapter 5 investigates how performance varies over time within the BCSP by following a large cohort of screening colonoscopists longitudinally to try to ascertain which factors predict the best performance and therefore highlight experts early.

Chapter 6 addresses the question: “What defines an expert endoscopist?” Semi-structured interviews with a randomly selected group of screening colonoscopists are analysed to identify themes that underpin true expertise in colonoscopy. What characteristics do the best endoscopists share?

Chapter 7 concludes the thesis, highlighting common themes from each of the studies and pointing towards future research possibilities.

2 The impact of the introduction of formalised polypectomy assessment in the United Kingdom

2.1 Background

Data from national audits[2,15] and from the English Bowel Cancer Screening Programme[20] have shown that most cases of bleeding and perforation are related to polypectomy. Training in polypectomy has, to date, been variable and poorly structured[84]. There has been some evidence suggesting poor exposure to polypectomy during training. In one study, only 60 per cent of trainees had ever been assessed in polypectomy[84].

2.1.1 JETS

As described in detail previously, in the United Kingdom application for formal certification of competency in colonoscopy has since October 2011 been via the electronic JETS website, having been a paper-based process before this date. Both colonoscopy and polypectomy assessments are now obligatory.

2.1.2 DOPyS

The Direct Observation of Polypectomy Skills assessment tool (Appendix 2), having been developed within the Bowel Cancer Screening Programme was introduced to trainees as part of the transition to electronic certification.

2.1.3 Minimum requirements

There have always been a set of minimum requirements before an application for certification of competency was permitted. Until October 2011, trainees had to show evidence of competency in colonoscopy by using the Directly Observed Procedural Skills (DOPS) tool for a summative assessment, using paper-based portfolios. At least four summative DOPS were required.

It was also a prerequisite that trainees had to have more than 200 completed colonoscopies. Polypectomy assessment itself was not compulsory, although trainers could reflect appropriate use of therapy during colonoscopy by marking a single criterion on the generic DOPS tool.

After October 2011, there was a transition to a paperless electronic certification system. Trainees had to complete the same summative DOPS assessments, as well as keeping a log of procedures. However, it was also compulsory for those wishing to be provisionally certified to have completed a minimum of four snare polypectomy assessments using the DOPyS to resect polyps up to 10mm as well as 10 formative colonoscopy DOPS during the course of their training.

2.2 Aims

2.2.1 Primary outcome

The primary aim was to describe the impact on documentation of polypectomy experience by the mandatory introduction of the DOPyS as part of the formal colonoscopy certification process.

2.2.2 Secondary outcomes

The secondary aims were to:

- evaluate how the introduction of the electronic certification system influenced trainees' exposure to colonoscopy during their training in relation to the mandatory minimum number of procedures
- assess if overall polypectomy performance could be predicted from the individual scores given in the DOPyS
- determine the strength of correlation between individual DOPyS descriptors and the final competency score.

2.3 Hypothesis

It is predicted that the introduction of an electronic portfolio will improve documentation of polypectomy and colonoscopy training.

2.4 Methods

2.4.1 Participants

All trainee endoscopists applying to JAG for certification of colonoscopy competence for the two years from October 2010 to October 2012 were included. The first cohort comprised those using paper portfolios from October 2010 to September 2011; the second cohort included those utilising the mandatory electronic portfolio from October 2011 to September 2012.

2.4.2 Data collection

Applicants to JAG for certification consent to their anonymised data being used for research. Permission for this audit was given by JAG central office. Applications for certification in the year prior to October 2011 were analysed retrospectively and compared with data collected prospectively for those in the following year.

2.4.2.1 Paper portfolios

Endoscopists were not required to submit procedural details for each case. Aggregate data were collected from the logbook on the total lifetime number of colonoscopies performed, the number of assessments submitted for both colonoscopy and polypectomy and whether applicants had any evidence of performing supervised or unsupervised polypectomy before certification of competence in colonoscopy.

If trainees using the paper-based certification system had any DOPS assessment marked demonstrating appropriate use of therapy, this was noted as experience of polypectomy. If

trainees had any documented experience of performing endoscopic mucosal resection (EMR), this was also recorded. Trainees' were allowed to use their own definitions of EMR.

2.4.2.2 Electronic portfolios

The prospective data collection parameters included numbers of formative DOPS colonoscopy assessments undertaken by each trainee, DOPyS assessments throughout the training period, total colonoscopies performed, total polypectomies, as well as data on the size, location and type of lesions resected by trainees when being assessed.

2.4.3 Statistical analysis

To assess differences between the two years of data collection, the Mann-Whitney U test was used for non-parametric data. The independent samples t test was used to compare normally distributed data such as polyp size.

Logistic regression was initially performed on the data obtained from the electronic DOPyS on a pass/fail basis with scores of 3 or 4 equating to a pass and 1 or 2 categorised as a fail. Ordinal regression was then used on the overall competency score of the DOPyS as a form of multivariate analysis.

Correlation between the 34 individual descriptors included in the DOPyS was performed using the Spearman rank correlation coefficient.

2.5 Results

2.5.1 Participants

There were 175 applicants for certification in the first year compared to 150 applications for provisional certification in the year (p=0.99) after DOPyS was introduced.

2.5.2 Experience of colonoscopy

Prior to seeking certification, the median number of colonoscopy procedures performed by each candidate was 287 in the first cohort, compared to 206 in the second cohort. There was no significant difference in the total number of colonoscopy procedures undertaken (p=0.07). These numbers were both above the requirement to complete at least 200 procedures before applying for certification.

In the first cohort, the median number of formative colonoscopy assessments provided was 3 (range 0-16). This rose to 32 (range 9-199) in the year after the introduction of the electronic portfolio (p<0.001, Figure 2.1), comprising the 10 mandatory DOPS required by the certification criteria and 22 additional DOPS submitted voluntarily.

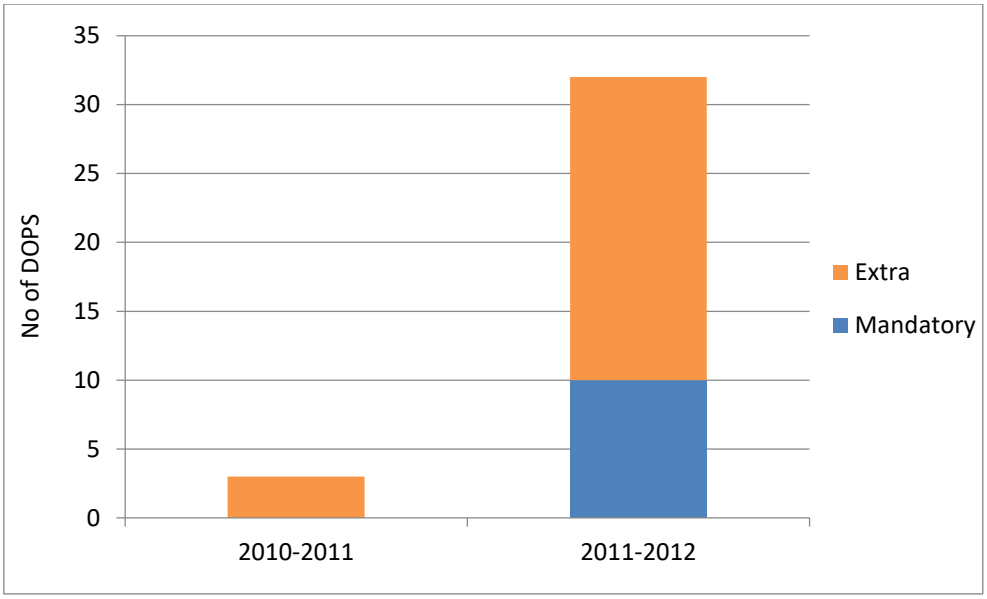

Figure 2.1. Formative colonoscopy DOPS submitted with applications for certification

2.5.3 Experience of polypectomy

Thirty two per cent of candidates using paper portfolios had evidence of any observed polypectomy in their logbooks with only 7 per cent of candidates referring to training in endoscopic mucosal resection (EMR).

In the year after the introduction of DOPyS, the 150 trainees utilising the electronic certification process were assessed resecting a total of 1283 polyps. All candidates had evidence of polypectomy assessment throughout that year with a median number of DOPyS of 7 (range 3-27). This figure included a median of 2 DOPyS for stalked lesions and 5 DOPyS for sessile polyps. Eighty nine per cent of applicants had evidence of assessed EMR on the liberal definition used in this research.

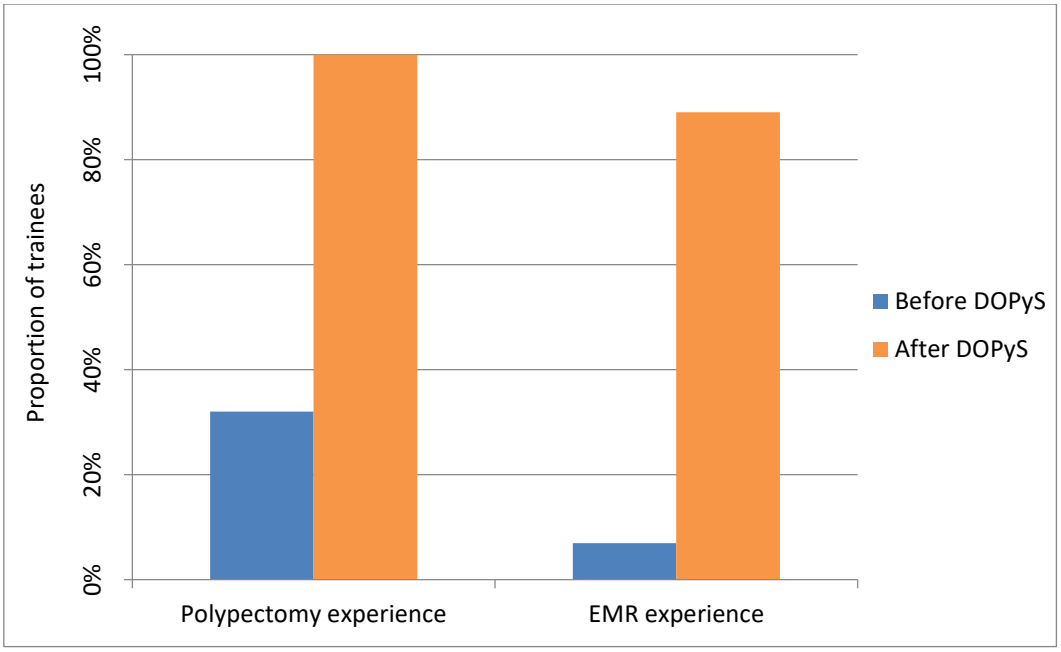

Figure 2.2. Proportion of trainees documenting any experience of polypectomy and EMR in their submissions to JAG

As expected, there was a significant increase in the number of recorded polypectomy assessments in the submitted data ($p<0.001$), given these were mandatory for the second cohort. However, logged experience of EMR, which was not obligatory also increased significantly ($p<0.001$).

2.5.3.1 Scoring of DOPyS

The majority of DOPyS assessments (93 per cent) were scored either 3 or 4 signifying competence at polypectomy (Figure 2.3).

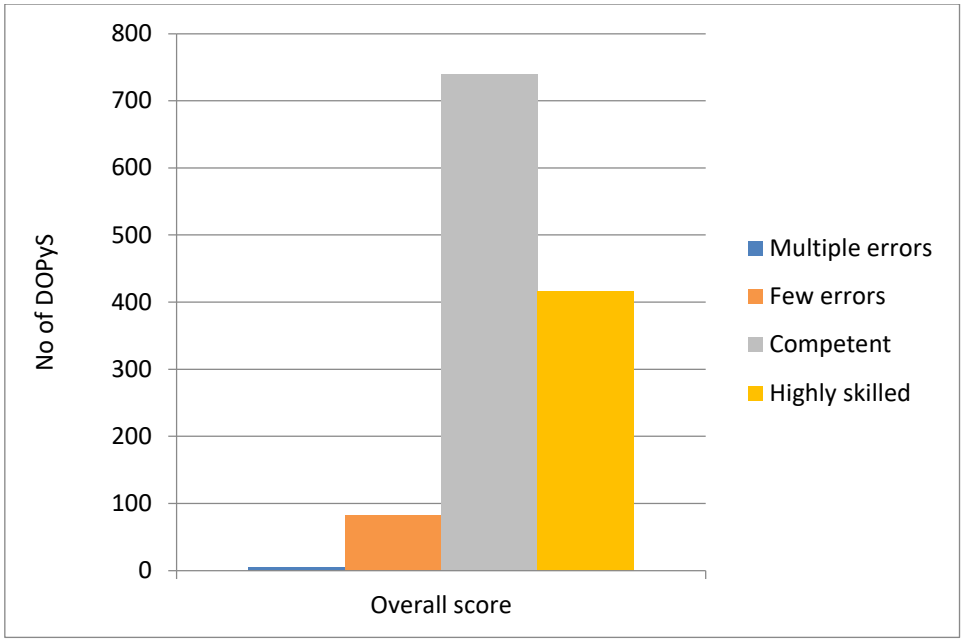

Figure 2.3. Overall polypectomy competency scored using the DOPyS

2.5.4 Types of lesions encountered

2.5.4.1 Location

Analysis of the DOPyS assessed polypectomies showed the majority of polyps (68.5%) were located in the left hemicolon, but a significant minority were found more proximally (Figure 2.4).

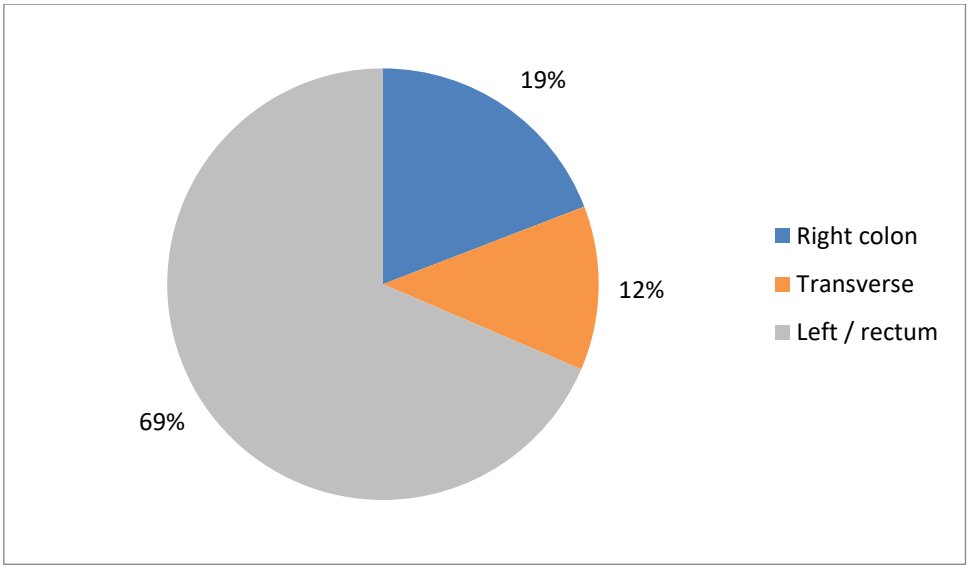

Figure 2.4. Location of polyps resected by trainees

2.5.4.2 Size and morphology

The polyps removed under assessment had a mean size of less than 10mm in all colonic segments. This was in keeping with expected standards as competency in polypectomy up to 10mm was required for provisional certification. These smaller lesions accounted for 79% of assessments. However, there was a significant range of polyp sizes seen, with lesions up to 60 mm documented (Figure 2.5).

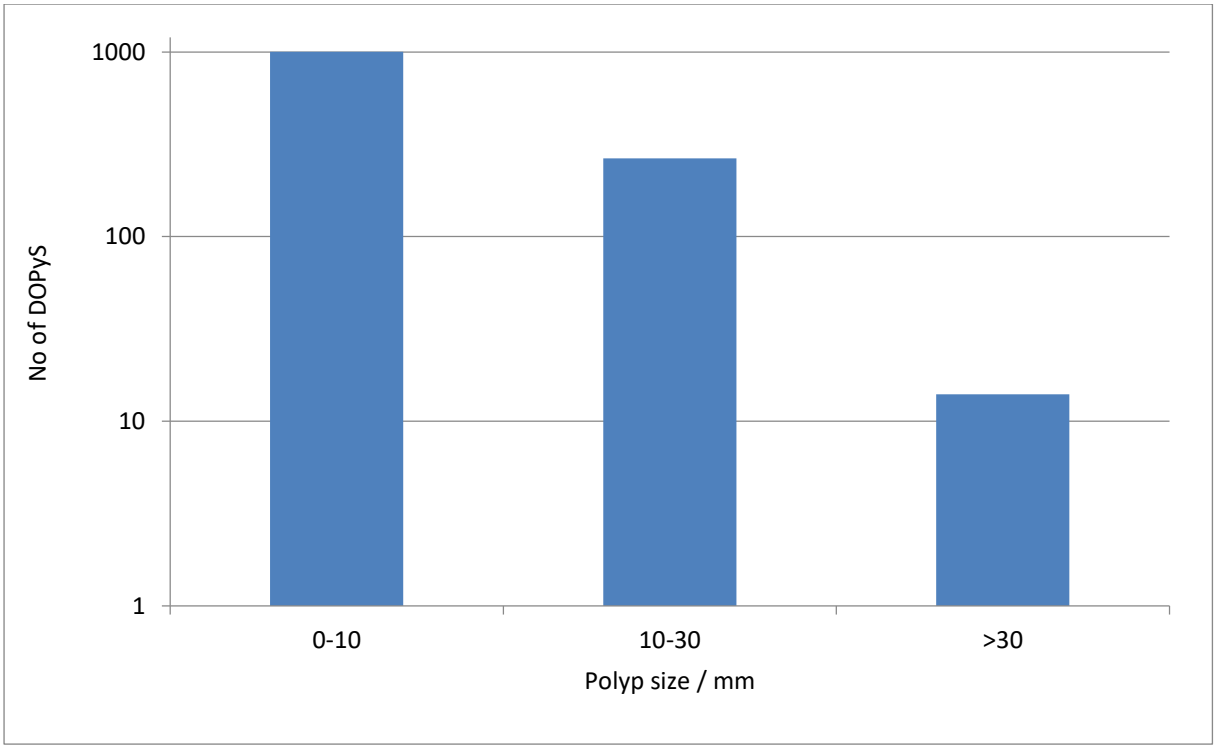

Figure 2.5. Polyp sizes assessed using the DOPyS

Polyps in the left colon were on average slightly larger (9.1 mm compared to 7.2mm, $p<0.001$). The sessile lesions tackled comprised the majority (61%) of polyps, but these were significantly smaller than the pedunculated polyps removed (6.6mm compared to 10.9mm, $p<0.001$).

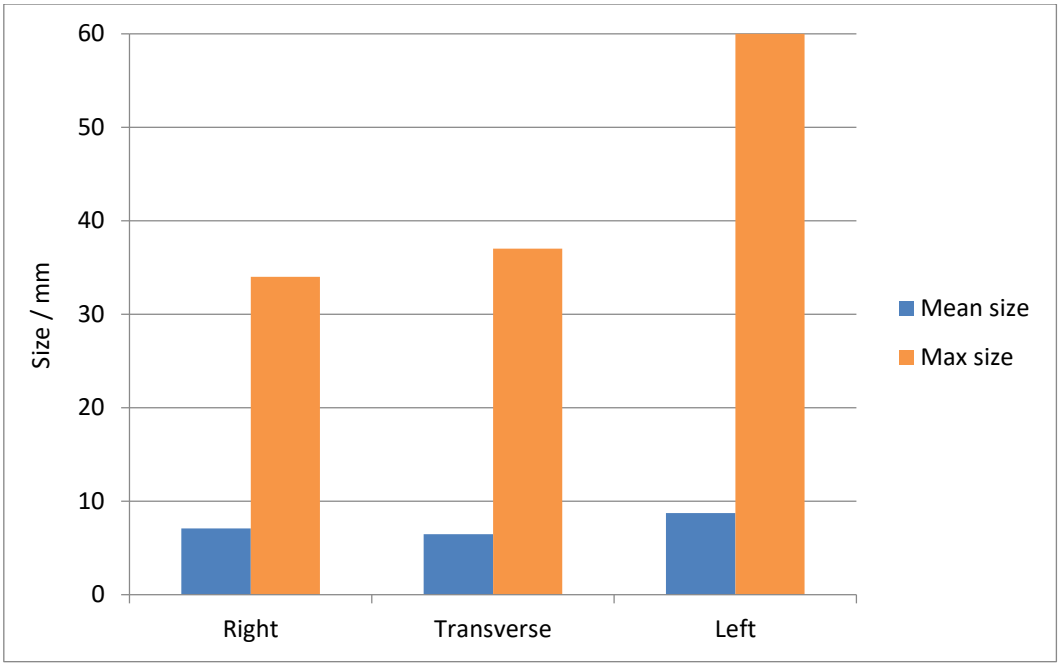

Figure 2.6. Mean and maximum polyp sizes removed by trainees with assessment

There was no difference in polyp size between the DOPyS scored as 1 or 2 compared to those scored 3 or 4 (8.0mm vs 8.3mm, p=0.299), implying that it was not the more difficult larger lesions that were scored less highly. However, it is possible that selection bias may have affected these results, with trainers only permitting competent trainees to tackle larger lesions.

2.5.4.3 Prediction of performance

Binomial logistic regression was used to try to generate factors that could predict overall polypectomy performance. When the outcome measure was a pass (3 or 4 on the DOPyS scoring system), polyp morphology, location or size did not affect the final score (p=0.692, 0.421 and 0.065 respectively).

However, when trying to predict the best performers, those attaining a score of 4, logistic regression did show a statistically significant association with a larger polyp size (p=<0.001), but did not demonstrate a difference by colonic location (p=0.236) i.e. those attempting larger polyps were more likely to score the best score on the DOPyS.

A model was then developed using ordinal regression on all the overall competency scores (range 1-4) attained during the DOPyS. This was statistically significant (p<0.001) with a good fit for the data (Pearson p=0.142). However, the effect of the tested variables (location, size, morphology) on the outcome was modest, as only a small amount of the variability in the final DOPyS score could be predicted from the data i.e. the individual components of the DOPyS could not easily be linked to the final score.

Table 2.1. Ordinal regression correlating polyp characteristics with DOPyS score

		Estimate	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Threshold	[Competency = 1]	-4.630	.475	**.000***	-5.560	-3.699
	[Competency = 2]	-1.667	.193	**.000***	-2.045	-1.290
	[Competency = 3]	1.738	.187	**.000***	1.372	2.103
Polyp location	Caecum	.524	.257	**.042***	.020	1.028
	Ascending colon	.262	.237	.269	-.203	.728
	Hepatic flexure	-.063	.427	.882	-.900	.773
	Transverse colon	.072	.222	.745	-.364	.508
	Splenic flexure	.373	.350	.287	-.313	1.060
	Descending colon	.148	.227	.514	-.296	.592
	Sigmoid colon	-.106	.177	.552	-.453	.242
	Rectum	baseline				
Morphology	Sessile polyp	.284	.133	**.033***	.024	.544
	Polyp size / mm	.032	.011	**.002***	.012	.053

The ordinal regression revealed that there were significant differences between each of the different DOPyS scores. Larger sessile polyps, those performed by endoscopists towards the end of their training and those in the caecum were more likely to be scored the highest marks (Table 2.1). The estimate column of data shows the effect of a one unit increase in the variable**,** for example a one millimetre increase in polyp size, in the ordered log odds of being in a higher level of DOPyS score.

A further analysis, where the individual DOPyS parameters were correlated with the final competency score, revealed significant correlations between all 34 parameters individually and also with the final score. A total of 630 separate correlation coefficients were calculated, all of which had a p value of less than 0.001. The mean Spearman's rank correlation coefficient was 0.41. The strongest correlation was between parameters "Ensures appropriate amount of tissue is trapped within snare" and "Tents lesion gently away from the mucosa" with a correlation coefficient of 0.836.

Logistic regression of the DOPyS parameters across the pass/fail scores revealed several parameters which were highly significant (Table 2.2). The B test statistic is the regression coefficient. In the model generated, this is positive for all the listed parameters, indicating that an increase in performance in these variables is linked to a higher overall DOPyS score. The Wald statistic is a chi-square analysis of the data testing whether the tested parameter, for example "Attempts to achieve optimal polyp position", makes a substantial contribution towards the overall score. Analysis of the Wald statistic gives the final p value and whether the parameter tested is statistically significant.

Table 2.2. DOPyS parameters most strongly associated with polypectomy competency

	B	S.E.	Wald	df	Sig.	Exp(B)
Attempts to achieve optimal polyp position	1.284	0.371	11.966	1	0.001	3.611
Adjusts / stabilises scope position	1.068	0.343	9.675	1	0.002	2.91
Checks all polypectomy equipment available	0.347	0.17	4.186	1	0.041	1.415
Clear instructions to and utilisation of endoscopy staff	1.021	0.244	17.482	1	0	2.775
Retrieves or attempts retrieval of polyp	0.372	0.136	7.518	1	0.006	1.451

2.6 Discussion

These data are the first to describe the implementation of a nationwide certification system on polypectomy training. The introduction of structured polypectomy assessment was both feasible and acceptable, with similar numbers of trainees applying for certification after the introduction of the DOPyS. The study shows a definite increase in the amount of documented supervised training overall received and clear evidence of an increase in the formal documentation of trainees' exposure to therapeutic endoscopy.

It is not possible to definitively say that actual training increased as a result of these changes as submission of information about trainees' polypectomies was not mandatory under the old system. There is some evidence in surgery that supervised training in technical skills does result in a higher rate of skill acquisition than unsupervised training[85]; similar data are not available for colonoscopic polypectomy.

The DOPyS was designed to aid both training and permit certification of competency. However it appears that it is currently being utilised predominantly with the latter objective: this may well be related to its inclusion as a mandatory part of the colonoscopy competency certification process. Only a very small proportion (0.4%) of assessments were scored 1 (accepted standards not yet met, frequent errors uncorrected). These low-scoring DOPyS are likely to represent formative assessments for trainees in the early stages of their learning curve.

The reasons for this are likely to be multifactorial. Given the hazards inherent to the technique, it is likely that a trainee performing a supervised polypectomy with frequent errors would either receive instruction from the trainer or pass the endoscope to the more experienced colonoscopist to complete the procedure. In addition, these evaluations are largely instigated at the behest of the trainee and it is likely that there is an element of selection bias with trainees not requesting trainers' formal documented assessment of procedures that have not been performed well. It is unlikely that the assessments are representative of the full learning curve for polypectomy given they are so heavily skewed towards exhibiting competency. These results are not unexpected, in that trainees were expected to use DOPyS to demonstrate their polypectomy skills if they wished to apply for certification and the data available are consistent with this approach by trainees.

The ordinal regression analysis lends credence to this view that DOPyS results were biased towards demonstrating competency. As a result of so few DOPyS assessments being scored 1 or 2, below the competency threshold, it was unlikely that the model could accurately predict the full gamut of factors associated with a more difficult polypectomy. In particular, it is well recognised that right colonic polypectomy is more challenging than that in the rectum where

access is much easier and it is surprising that the regression did not show any influence of the polyp location on polypectomy difficulty.

Interestingly, the individual DOPyS parameters most associated with polypectomy competency included both technical and non-technical skills, including descriptors such as "clear instructions to and utilisation of endoscopy staff". It may be that descriptors such as this are associated with seniority, with more junior endoscopists feeling less confident in asking other staff for help.

It is encouraging that all trainees are receiving dedicated separate assessment in both colonoscopy and polypectomy. Just over 10 years ago, the UK national audit showed that only 17.0% of colonoscopists had received supervised training for their first 100 colonoscopies (section 1.4.1.1)[15]. This study shows the significant progress made in the assessment and documentation of training in a relatively short space of time.

It is also reassuring that trainees are being taught and assessed polypectomy on a variety of lesions, including the more risky and technically challenging proximal polyps. Although most assessed lesions were less than 10mm in size, in keeping with the parameters of the mandatory DOPyS assessment, this mirrors clinical practice. In the FOBT positive Bowel Cancer Screening Programme, 86% of polyps were found to be less than 10 mm and it seems that training experience reflects these figures[86].

The vast majority of polyps found during universal screening flexible sigmoidoscopy at age 55 are also within the same size parameters, emphasising the importance of all endoscopists possessing the skills required to tackle these common lesions. In this screening setting,

complete polypectomy and retrieval is crucial. Although larger sized lesions are associated with higher rates of advanced pathology such as high-grade dysplasia or villous architecture, these features are not confined to bigger polyps and occur in lesions less than 5mm[87]. It must be noted that one of the limitations of the DOPyS is that it does not distinguish between sessile and flat polyps; these lesions require differing endoscopic approaches and at present it is not possible to assess resection of a flat lesion using this tool.

Only a very small minority (7%) of trainees before the introduction of DOPyS had evidence of endoscopic mucosal resection skills. The methodology used to assess this figure was generous with even a solitary reference to EMR being counted. These data were obtained from logbooks and so the quality and amount of training could not be easily gleaned.

However, since DOPyS has found its way into routine clinical practice with trainees performing more assessments than strictly mandated, this would suggest that the majority of trainees seem to be receiving both training and assessment in this invaluable technique, a skill which many would argue is obligatory for any independent colonoscopy practitioner.

Interestingly, with the introduction of the electronic portfolio, most trainees considerably surpassed the minimum requirements for submission of formative DOPS and DOPyS. This suggests that trainees have become familiar with this method of assessment and integrated it into their practice, using it as a training aid as well as for certification of competency. As polypectomy plays an increasing role globally in colorectal cancer prevention, the DOPyS provides an effective means of assessing and certifying polypectomy in order to minimise the well-recognised risks associated with this technique.

The role of assessment in changing academic behaviour has been described before (section 1.5.1.3). The closest parallel with practical medical skills assessment comes from the DOPS tool, which is used widely by junior doctors as part of their electronic portfolios. A study looking at doctors who had just qualified from medical school found that most (70%) felt that DOPS helped improve their clinical skills[88]. However, a systematic review of several different types of workplace-based assessments found no available high-quality evidence from this study or any others that DOPS lead to objective performance improvement[89].

The LAPCO program, a training structure devised to encourage the widespread uptake of laparoscopic colorectal surgery, has been extensively evaluated[90] and bears some similarity to polypectomy assessment. Similar to the DOPyS, task deconstruction, hands-on training and independent expert rating of cases with feedback are hallmarks of this program[91]. Analysis of outcome data has shown significant improvement in participating surgeons, with expert levels of proficiency achieved after iterative feedback[92].

The best studied tool with respect to performance improvement is the multisource feedback, in which both senior and junior colleagues anonymously provide ratings of an individual's strengths and weaknesses. The ability of this to change behaviour is reliant on individual characteristics. Studies have shown that some junior doctors[93] and most surgeons[94] were not amenable to change whereas general practitioners were more likely to be receptive to suggestions[95]. Those who consider change to be necessary, react positively, believe that changes feasible and take suitable actions as a result of this are most likely to benefit from feedback[89].

It is likely that the DOPyS would show similar characteristics, but this has not been evaluated to date. Specifically, the DOPyS focuses predominantly on technical skills involved in polypectomy. These are easier to influence than non-technical skills which require behavioural change and insight. A meta-analysis of non-technical skills training revealed that outcome measures utilised to date in the literature are seldom clinically applicable[96]. However, the data described here does show some correlation with non-technical skills; studies published to date in other fields may not be applicable to polypectomy.

Several questions remain unanswered from this work. Whereas there is growing consensus in the literature as to the number of colonoscopies needed to attain the technical ability to reach the caecum reliably[46,50,63], this data is lacking for polypectomy skills. Training in polypectomy is inherently high risk and the quality of training and assessment is likely to vary considerably between specialties, centres and individual trainers. The relative importance of non-technical skills when undertaking polypectomy at an introductory level is also not clear but they are likely to be important. Assessment of these other skills is likely to be more difficult than simple technical ability but may prove to be beneficial in terms of improving the rate of skill acquisition.

This study is an important first step in addressing some of these issues. The next chapter will address the issues concerning variability in training between countries specifically and future work will aim to formally establish the learning curve in polypectomy.

3 An international survey of polypectomy training and assessment

3.1 Background

The use of colonoscopy for both diagnosis and screening is increasing worldwide. Variability in colonoscopist performance is well-documented. In an effort to ensure that practitioners achieve competence before working independently, many countries have been establishing competency-based assessments of colonoscopy technique, as opposed to requiring purely numerical documentation of completed procedures.

The most hazardous part of colonoscopy is polypectomy, accounting for the majority of serious complications: post-colonoscopy bleeding and perforation. Training in polypectomy has been variable with some trainees receiving little in the way of formal guidance before being expected to perform colonoscopy with polypectomy independently. In addition, until recently there has been no validated structured way to assess competency at polypectomy.

The Directly Observed Polypectomy Skills (DOPyS) tool has been devised to permit documentation of polypectomy competency. In the United Kingdom, polypectomy assessment has now been deigned mandatory as part of the colonoscopy certification process, with trainees having to demonstrate competency in the removal of polyps up to 1cm for provisional certification and between 1 and 2 cm for full certification.

It is unclear whether other countries have highlighted polypectomy as a specific skill that needs to be taught. There is also no data on how training and assessment of polypectomy are currently delivered around the world.

3.2 Aims

3.2.1 Primary outcomes

3.2.1.1 Assessment of polypectomy training in clinical practice

The first aim of the study was to assess both trainees' and trainers' experience of polypectomy training in countries around the world.

3.2.1.2 Assessment of guidelines for training

The second aim was to ascertain whether guidelines exist for polypectomy competency assessment in countries apart from the UK.

3.3 Hypothesis

It is hypothesised that only a minority of trainees will have experienced structured polypectomy training and assessment.

3.4 Methods

3.4.1 Participants

Participants were recruited in several ways. International leaders in the field of colonoscopy, many of whom belong to World Endoscopy Organisation (WEO), were initially asked if they were willing to participate in an international survey. If there was no representative identified from the WEO, or the endoscopist declined to participate, an email was sent to an

endoscopist known to be interested in training asking if they wished to be involved in the project.

3.4.2 Survey

Local representatives in each country were asked to provide email access to a national database of trainees or, where this was not available, a regional list of endoscopy trainees comprising where possible both those with a medical and surgical background. A further list of endoscopy trainers was also requested.

3.4.3 Guidelines

Each representative was asked to forward guidelines from their respective countries relating to polypectomy training and competency assessment. If no guidelines existed, members were asked to confirm this.

3.4.4 Survey

An online survey was created asking separate questions to trainees and trainers. The survey was administered in English. The information sought is shown in Table 3.1.

Table 3.1. Information collected from endoscopists

Trainee	Trainer
Specialty	Specialty
Awareness of national guidelines in polypectomy	Awareness of national guidelines in polypectomy
Colonoscopy experience	Use of guidelines when training
Formal polypectomy teaching	Frequency of delivering training
Polypectomy experience	History of polypectomy assessment training

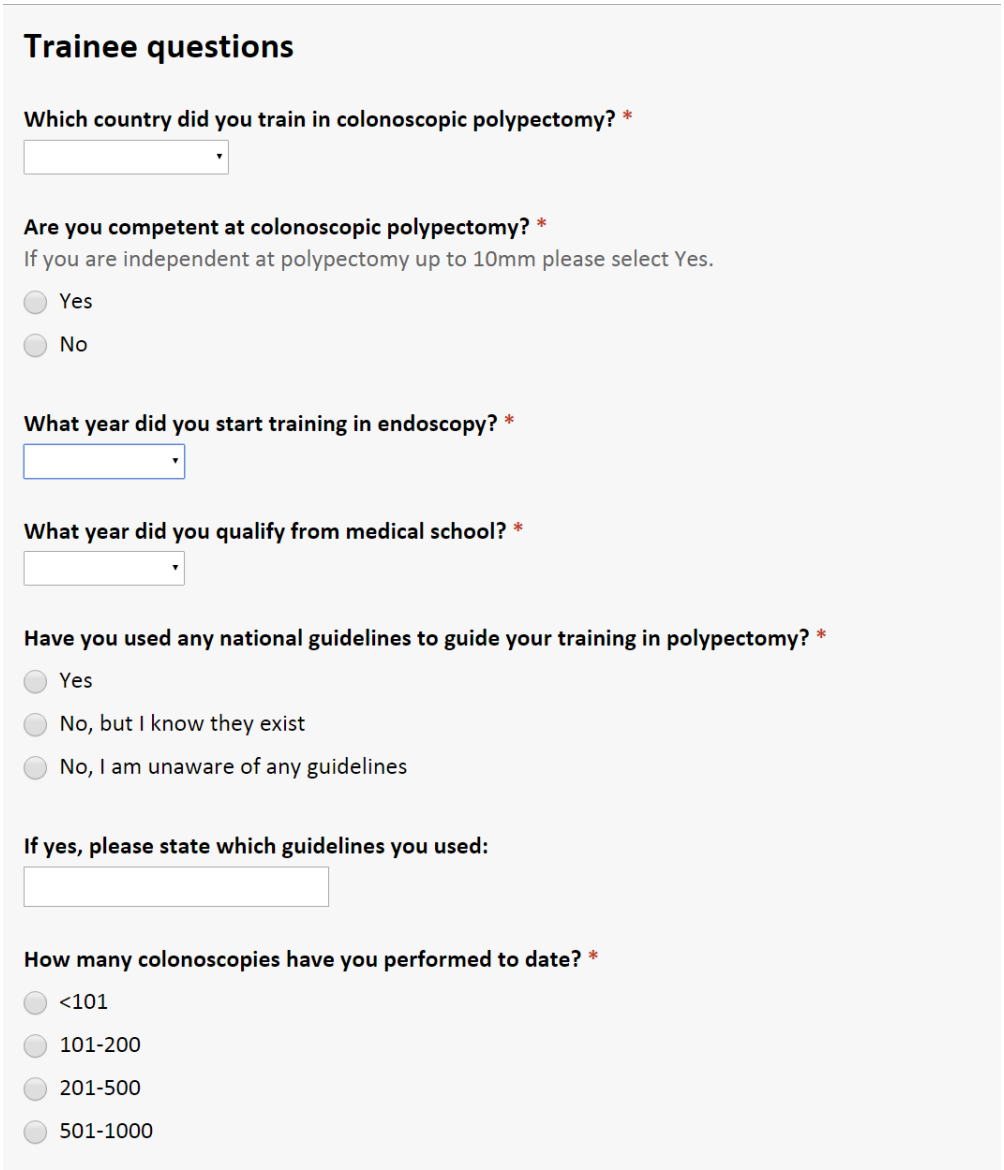

Figure 3.1. Screenshot of part of trainee survey

3.5 Results

3.5.1 Participants

3.5.1.1 Demographics

Data were obtained from 610 colonoscopists. Of these responses, 348 (57.0%) were from trainers and 262 (43.0%) from trainees.

3.5.1.2 Country of origin

In total, 19 countries or territories were represented in the survey spanning five continents (Table 3.2 / Figure 3.2).

Table 3.2. Countries or territories participating in the study

Australia	**Hong Kong**	**Spain**
Austria	Hungary	Thailand
Canada	Japan	Italy
France	New Zealand	United Kingdom
Germany	Nigeria	United States of America
Greece	Sudan	
Holland	Poland	

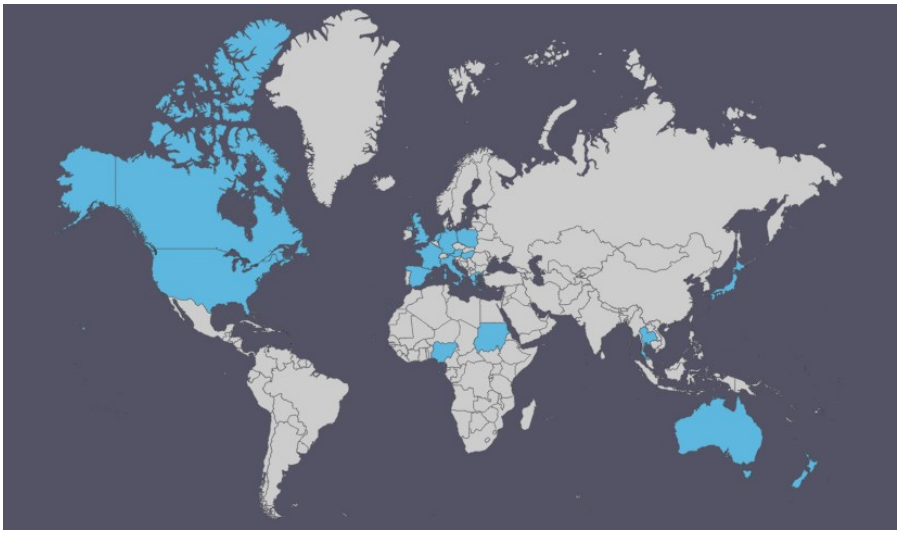

Figure 3.2. Countries or territories participating in the study

3.5.1.3 Trainers

3.5.1.3.1 Origin

Trainers originated from all 19 countries participating in the survey (Figure 3.3). The biggest contribution to this part of the survey came from Spain, with 86 responses recorded.

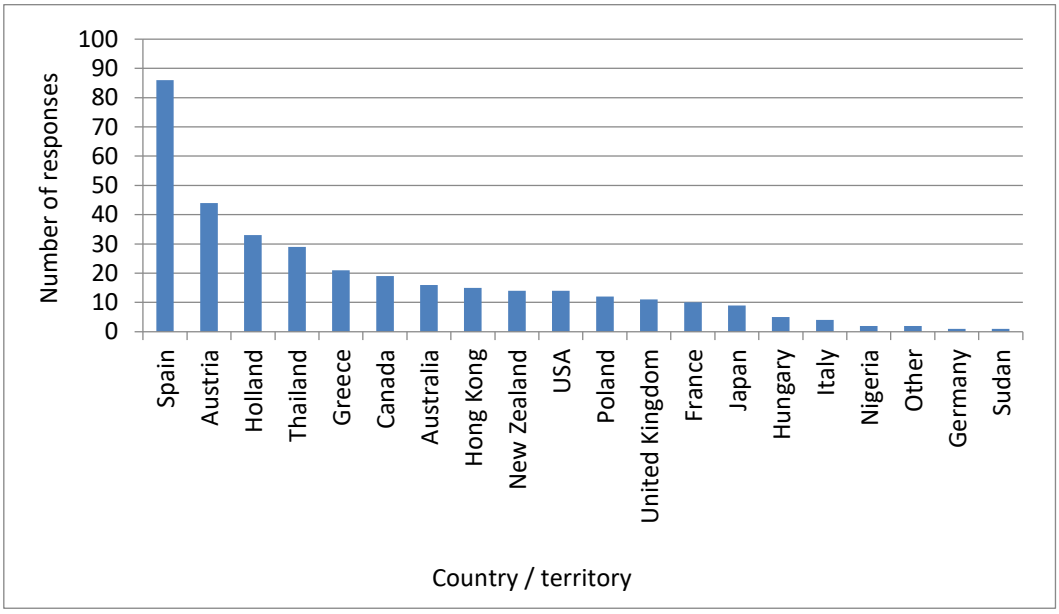

Figure 3.3. Trainer responses by country

3.5.1.3.2 Background

The majority (92.2%) of trainers had their origins in medical gastroenterology, with a 6.6% (23/348) having a background in surgery.

3.5.1.4 Trainees

3.5.1.4.1 Origin

Trainees from 17 countries responded to the survey. The two countries from which no trainee responses were obtained were Germany and Nigeria.

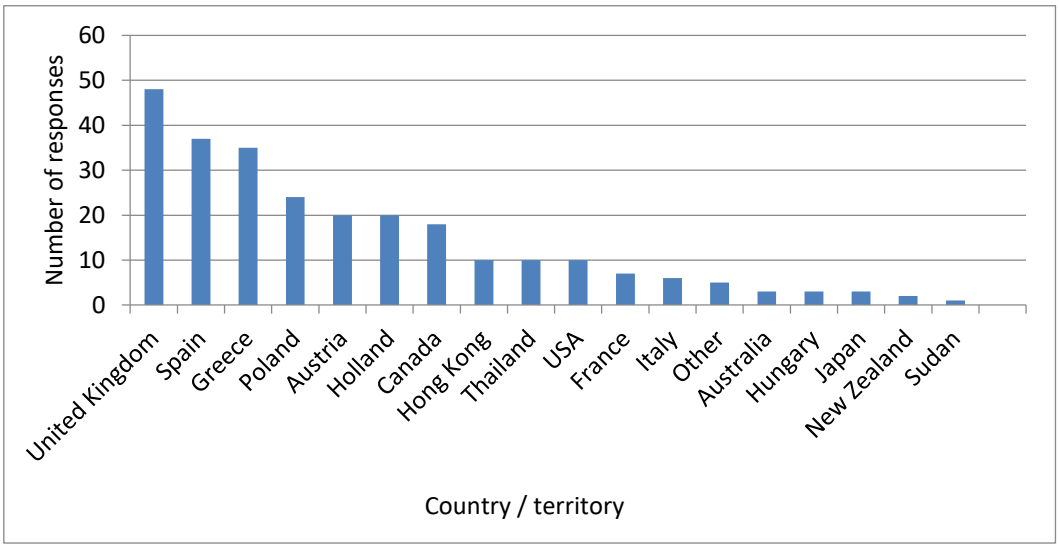

Figure 3.4. Trainee responses by country

3.5.1.4.2 Background

Again, the majority (90.5%, 237 respondents) of trainees had a primary specialty of medical gastroenterology, with a small proportion (3.8%, 10 respondents) training in surgery.

3.5.2 Trainer survey

3.5.2.1 Frequency of assessment

The majority (79.6%) of the 348 trainers surveyed were involved in trainee polypectomy assessment weekly. Twelve respondents considered themselves polypectomy trainers but never formally assessed their trainees' competency at polypectomy and 11 trainers stated they only performed assessments annually.

3.5.2.2 Awareness & utilisation of guidelines

A small majority of trainers (58.9%, 205 respondents) were aware of the existence of guidelines about training in polypectomy. Of this group, 180 (87.8%) stated that they used polypectomy guidelines when training endoscopists, with 25 (12.2%) not using any guidelines.

Multiple responses were received from trainers when asked to state the guidelines that they used. The most commonly cited guidelines, mentioned by fifty respondents were the ASGE guidelines. Significant numbers also said that they used BSG, AGA and European guidelines (Table 3.3).

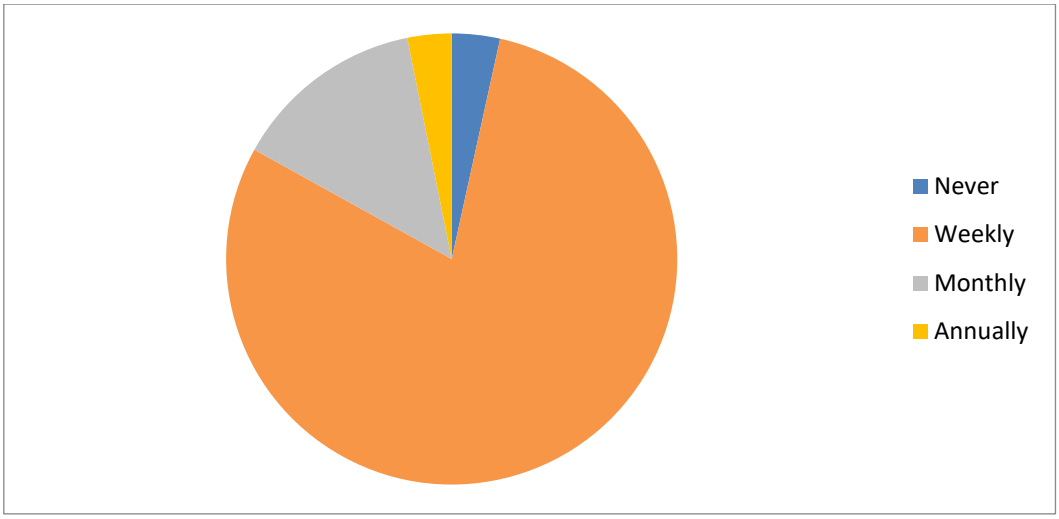

Figure 3.5. Frequency of polypectomy competency assessments by trainers

Table 3.3. Guidelines quoted by trainers to assess polypectomy

Guideline	Number of responses
ASGE	50
European / ESGE	33
AGA	19
BSG	16
DOPyS	7
In house / local	13

3.5.2.3 Awareness of framework for assessing polypectomy

Just over half (51.4%, 178 respondents) of the trainers surveyed said that they used a framework when assessing polypectomy. Most of these individuals (130 respondents) were those who used polypectomy guidelines to train other endoscopists.

3.5.3 Trainee survey

3.5.3.1 Colonoscopy experience

Most trainees (76.7%) responding to the survey commenced endoscopic training within the years 2008-2012 and qualified from medical school after the year 2000 (80.9%). The survey included those with a breadth of colonoscopic experience (Figure 3.6) with 31.7% having completed more than 500 colonoscopies and 38.2% having completed fewer than 200 procedures.

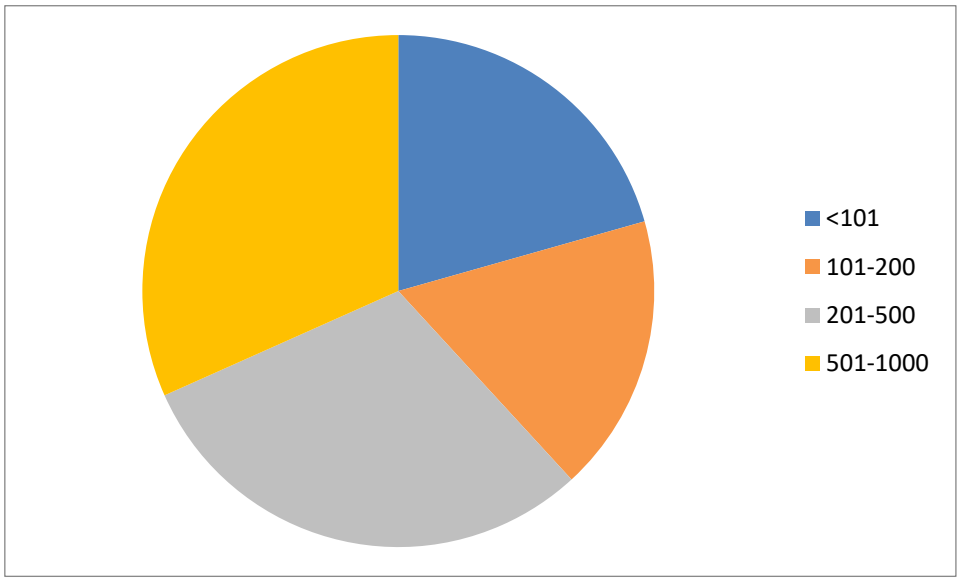

Figure 3.6. Number of colonoscopies performed by trainees

3.5.3.2 Polypectomy experience

3.5.3.2.1 Theoretical

Trainees were asked whether they had been formally taught the principles behind polypectomy. Generally training improved as experience increased. A minority had never received such teaching, with approximately half (51.1%) stating that these principles had only

been taught intermittently. A significant minority (8.4%) of those who had performed the most colonoscopy had never been taught about polypectomy (see Figure 3.7).

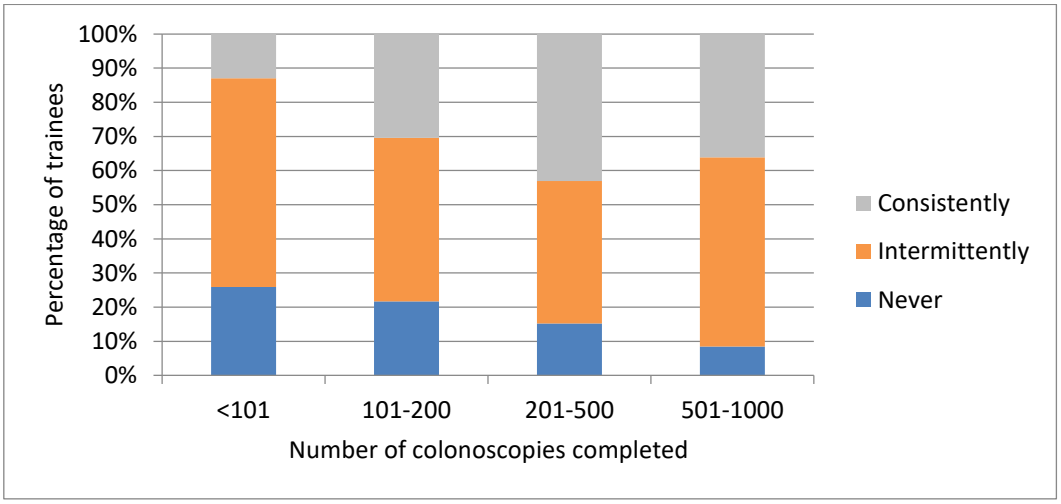

Figure 3.7. Number of colonoscopies performed by trainees and proportion taught the principles of polypectomy

Most (64.1%, 168 respondents) trainees had never been taught the principles of endoscopic mucosal resection (EMR). Again, the proportion of those training in EMR principles increased with colonoscopic experience, although even among those with significant experience, having performed more than 500 colonoscopies, only 49.4% per cent had received teaching on the theory behind EMR (Figure 3.8).

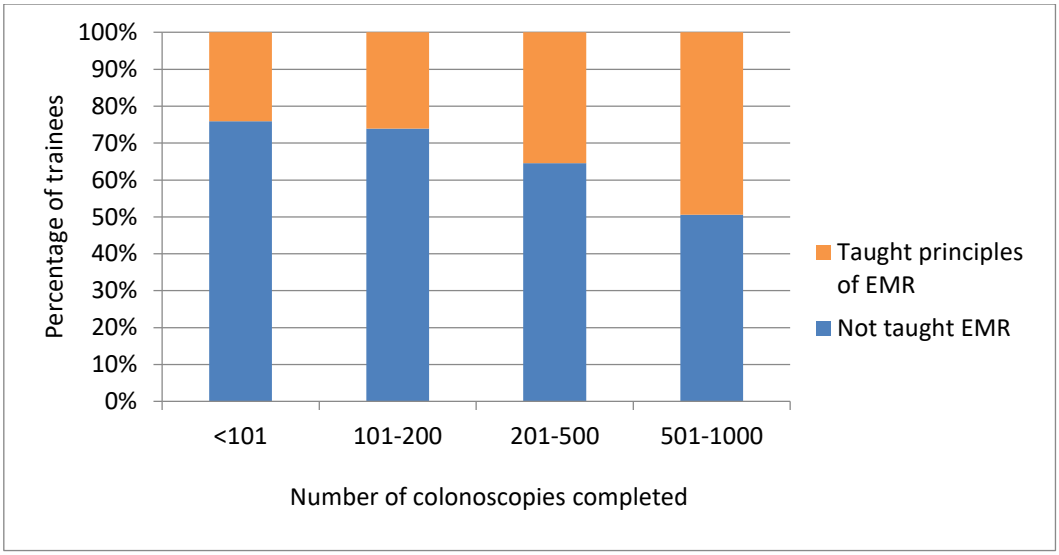

Figure 3.8. Number of colonoscopies performed by trainees and proportion taught the principles of EMR

3.5.3.2.2 Use of guidelines

A minority of trainees used guidelines to direct their polypectomy training. In total, 32.8% used national guidelines but a larger proportion, 43.1%, were aware of the existence of guidelines but chose not to use them. Seventeen trainees quoted American guidelines, encompassing the AGA and ASGE, with all 22 trainees from the United Kingdom who stated they used guidelines when training utilising JAG guidance. No trainees from outside the United Kingdom mentioned JAG or DOPyS, instead usually referring to their own national or continental guidelines.

3.5.3.3 Hands-on training

3.5.3.3.1 Assessment

Just over half (53.1%, 139 respondents) of trainees had ever had their polypectomy technique formally assessed by any trainer. Of the 262 trainees, 67.6% (177 respondents) stated that

they were competent at polypectomy. Of this self-certified competent group, 39.5% (70 respondents) of colonoscopists had never had a formal evaluation of their polypectomy technique.

3.5.3.3.2 Documentation

3.5.3.3.2.1 Polypectomy assessment

When asked how polypectomies that had been formally assessed were documented, 52.3% of trainees (68/130) used either a formal polypectomy assessment form or other generic form. However, 49.2% (64/130) stated that the polypectomy assessment process was not documented in any form.

3.5.3.3.2.2 Logbook

Trainees were also asked about whether they kept record of polypectomies performed in a logbook. 46.1% of trainees held no record of the polypectomies they had performed; 43.1% however did so regularly.

3.5.3.3.2.3 Large polyps

Most trainees had not received training for larger polyps, those over 10 mm. Overall, only 32.8% said that they had been taught specific skills to deal with these larger lesions. Of those with the most colonoscopy experience, who had performed over 500 procedures, 51.8% had had large polypectomy instruction.

Trainees who deemed themselves competent had had similar levels of training for larger polyps with 42.9% having been trained specifically in this way.

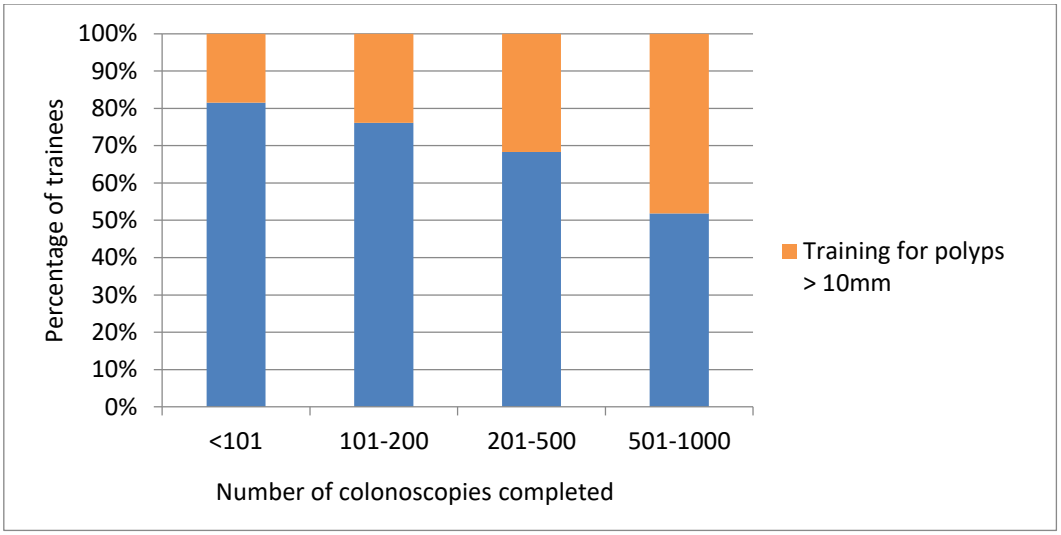

Figure 3.9. Proportion of trainees who had received specific training for larger polyps by total colonoscopy experience

3.5.4 Guidelines review

3.5.4.1 Colonoscopy

Training requirements for colonoscopy certification of competency in Europe and the USA have been described in detail in the introduction.

3.5.4.2 Polypectomy

Responses were received from each of the local representatives outlining polypectomy training in their own country. Only a minority of countries had guidelines devoted to polypectomy (see Table 3.4).

Table 3.4. Polypectomy training requirements in countries surveyed

Country	Polypectomy guidance
Australia	30 snare polypectomies needed
Austria	No polypectomy guidance
Canada	No polypectomy guidance
France	No polypectomy guidance
Germany	No polypectomy guidance
Greece	No polypectomy guidance
Holland	No polypectomy guidance
Hong Kong	No polypectomy guidance
Hungary	No polypectomy guidance
Japan	No polypectomy guidance
New Zealand	No polypectomy guidance
Nigeria	No polypectomy guidance
Sudan	No polypectomy guidance
Poland	No polypectomy guidance
Spain	No polypectomy guidance
Thailand	No polypectomy guidance
Italy	No polypectomy guidance
United Kingdom	JAG guidance – numbers and technique via assessments with DOPyS. See section 1.5.1.3.1.1
United States of America	No polypectomy guidance

3.5.5 Discussion

The majority of countries have no specific guidance on how polypectomy competency should be documented or how trainers should undertake assessing trainees. Even in countries where such guidance does exist, such as the United Kingdom, a significant proportion of trainees

were unaware of such guidance or did not use it in their training. Both of these issues need addressing.

There have been significant recent advances in endoscopic knowledge and techniques and it seems prudent that trainees benefit from this to improve patient outcomes. The demise of hot biopsy as a method for removing diminutive polyps due to the increased risk of perforation is a good example of the importance of ensuring trainees are instilled with both the appropriate technical and judgement skills necessary to perform polypectomy safely. In an era where the utilisation of polypectomy as a tool to prevent colorectal cancer in often healthy individuals is increasing inexorably, the gaps in training identified by the survey are concerning.

The majority of trainers in the survey were aware of what they considered to be national guidelines regarding polypectomy assessment. Nevertheless, analysis of these guidelines demonstrated little reference to the mechanics of polypectomy assessment and training, instead largely referring to documents detailing surveillance intervals for patients with polyps.

In addition, most trainers did not have a structured way in which to assess polypectomy, which makes the reproducibility of both teaching and the assessment of competency of trainees challenging, especially when several trainers have differing approaches.

Only a minority of trainers had received training themselves on how to assess competency at polypectomy. Previous work has shown that there is significant variability in both polypectomy technique and assessment of competence. With complex, multi-step tasks such as polypectomy, a standardised approach across trainers and different centres is likely to benefit trainees and improve the rate of skills acquisition.

It is disappointing that internationally a significant number of trainees were aware of the existence of national guidelines specifically relating to polypectomy but did not use them in their own training. There are likely to be myriad reasons for this including individual motivation, perceived lack of relevancy, and the attitude of trainers.

The lack of specific training in EMR and larger polypectomy is also a concern. Many trainees in the survey had completed what would be regarded as very significant numbers of colonoscopies, yet had never been taught or assessed on polypectomy. The concern is that some endoscopists may be deemed competent at therapeutic colonoscopy including polypectomy after attaining high performance metrics at diagnostic colonoscopy only. These colonoscopists may be exposing patients undergoing polypectomy to an excess risk.

In the United Kingdom the differences between training in colonoscopy and polypectomy are addressed by having a two tier certification system whereby trainees are first judged competent at diagnostic colonoscopy and sub-centimetre polypectomy and only then can progress to performing supervised polypectomy on larger lesions with full certification thereafter.

The vast majority of publications relating to competency in colonoscopy relate to the skills required to attain the technical ability to reach the caecum reliably. It is only recently that a consensus has developed on the number of colonoscopies that trainees need to perform to be able to pass the endoscope safely to the caecum[50].

The reasons for this focus are likely to be twofold. First, in the past diagnostic colonoscopy skills were often poor with significant numbers of patients undergoing incomplete

colonoscopy[15]. Diagnostic standards have risen in many countries but it seems that training in therapeutic colonoscopy needs to undergo a similar transformation.

Secondly, the measurement of the caecal intubation rate is a simple binary outcome for each patient. As such, this lends itself to easy measurement and analysis. By contrast, polypectomy is only performed in a minority of patients and the skills involved differ significantly between patients due to the inherent complexities of the technique.

In addition, whereas the caecal intubation rate is widely accepted as a quality mark for colonoscopy, no such consensus exists for the quality of polypectomy. The DOPyS has endeavoured to bridge this gap and has gained some international acceptance but this is by no means universal. It is clear however that competency in diagnostic colonoscopy does not confer similar competency in therapeutics[52]. The introduction of DOPyS into Bowel Cancer Screening accreditation was in direct response to the initial process testing diagnostic colonoscopy skills to a high standard but neglecting the key mandatory therapeutic aspects of the procedure.

In many ways, the literature for skill acquisition in more advanced techniques such as endoscopic submucosal dissection (ESD) is more mature than that of standard polypectomy. Several publications have addressed the challenges of the learning curve with respect to ESD[97-101]; a PubMed search of the terms “ESD learning curve” reveals over 40 pertinent publications. However, a similar search looking for “polypectomy learning curve” yields no relevant results despite the fact that polypectomy is performed far more widely and hence is responsible for considerably more morbidity than ESD.

The reasons for this are complex and may include that ESD is a new technique and that variability in outcomes has been definitively shown according to experience[102,103]. Similar data are lacking for endoscopic polypectomy.

Data from the English Bowel Cancer Screening Programme analysing procedures carried out over a five-year period examined over 167000 polypectomies with 850 patients experiencing bleeding[20]. In contrast, ESD is only performed on select patients in a few centres and by virtue of this is likely to cause significantly less morbidity than standard polypectomy.

Given that the usage of colonoscopy is increasing year on year worldwide, the amount of therapy performed and the proportion of procedures deemed therapeutic is also correspondingly increasing. Nevertheless, only a minority of the population in most countries undergo colonoscopy. Non-invasive technologies which are ancillary to colonoscopy at present are likely to eventually supersede the role of colonoscopy for the diagnosis of colonic polyps and consequently lead to a further increase in therapeutic procedures. This effect has already been observed with the increasing utilisation of CT colonography.

A similar shift has already occurred in an allied endoscopic field: endoscopic retrograde cholangiopancreatography (ERCP). In the past, this was performed for myriad indications in order to try to obtain a diagnosis in patients with suspected biliary pathology. However, the advent of high quality non-invasive imaging has now rendered this an exclusively therapeutic technique. A similar transition is likely to occur over time in respect of colonoscopy. Indeed, the purely diagnostic colonoscopist, unable to perform polypectomy, should no longer exist given the integral nature of polypectomy to colonoscopy in current practice.

There is much current endoscopic research on the removal of large benign and early malignant colonic lesions, which in the past would only have been deemed curable by formal operative resection. Rapid advances in fields such as ESD have increased the gamut of the endoscopic resection, but at the expense of requiring a highly skilled operator. The first step in training in these advanced therapeutic techniques is of course basic polypectomy, where skills such as fine tip control are mandatory for a successful and safe procedure.

In conclusion, this survey shows there is huge variation in polypectomy training and few formal guidelines. Despite this, there is no published evidence of poorer polypectomy outcomes between countries with different training regimens for their endoscopists. Capturing these differences however, with outcome metrics that are clinically meaningful such as interval cancer rate, is likely to be challenging given the different clinical environments around the world.

At present, it seems that most training worldwide is informal and without structure, in the way of much medical training in the past. However, with demands increasing from both professional regulators and patients for quantifiable medical competence, a more formalised approach than has been traditionally used is likely to be needed to address concerns not just about actual competence but evidence of competency.

There is a need for an international consensus as to what constitutes a competent polypectomy. In addition, the learning curve for this procedure needs to be defined to provide those learning therapeutic colonoscopy guidance so the many successes of colonoscopy in preventing colorectal cancer are not marred by iatrogenic complications. Finally, it would

seem prudent for those delivering training to be trained themselves to ensure standardisation in the methods they used to teach and assess trainees in this complex skill.

3.5.5.1 Strengths/limitations

This is the broadest geographical survey to have ever been conducted, specifically assessing the international experience of polypectomy from countries around the world. The inclusion of both trainees and trainers from geographically disparate areas is a strength. However, the method of recruitment of those to be included, partly through the expert forum of the WEO and partly through personal contacts interested in training in endoscopy, is likely to have resulted in selection bias with endoscopists at higher volume, academic centres more likely to have been represented. This bias is most likely to have been reflected by these results showing the best case scenario: it is likely that the overall training landscape in the country surveyed is poorer than that represented in the study.

In many cases, it was not possible to survey the entire nation's endoscopic trainees as many countries do not keep reliable central databases of those undergoing endoscopic training.

4 Screening colonoscopists' performance after a structured accreditation process in the English Bowel Cancer Screening Programme

4.1 Background

4.1.1 Inception of BCSP

A strong evidence base for colorectal cancer screening combined with poor mortality data in United Kingdom, when compared to Europe and the USA, were strong drivers towards the introduction of the Bowel Cancer Screening Programme in England.

The poor colonoscopy performance seen in a national audit in United Kingdom has been described previously. To ensure those entering a screening programme were of a high enough standard, a multi-stage accreditation process (section 1.5.1.3.1.2) was devised by experts.

Before the development of this process however, an informal colonoscopy skills assessment was conducted with invited colonoscopy experts. These individuals would make up the first cohort of the screening programme. In time, further colonoscopists were recruited using the DOPS and DOPyS tools (Appendices 1 and 2) as gauges of quality.

4.2 Aims

4.2.1 Primary outcome

The primary outcome was to ascertain whether data from the accreditation process correlated with meaningful clinical outcomes in the form of key performance indicators.

4.2.2 Secondary outcome

The secondary outcome was to determine how colonoscopy performance varied in this highly selected group of accredited colonoscopists in the 12 months after they commenced screening.

4.3 Hypothesis

It is predicted that the accreditation will be able to differentiate between the highest performers and the competent endoscopist.

4.4 Methods

4.4.1 Approvals

Endoscopists in the Bowel Cancer Screening Programme give their consent to their anonymised data being analysed for research purposes. The formal study proposal for an observational study was peer reviewed by the Bowel Cancer Screening Programme Research Committee, who gave formal permission for the study.

4.4.2 Participants

Data were collected up to March 2012 on all 244 screening colonoscopists who had undertaken 12 months of screening activity.

4.4.3 Data collection

Every procedure performed under the BCSP is recorded on an online database contemporaneously by an independent specialist screening nurse practitioner (SSP). The information recorded includes patient demographics, withdrawal time, anatomical extent of procedure, polyps detected and details of their management, drugs administered and patient comfort.

Every colonoscopy has a patient comfort score recorded by the SSP, without input from the endoscopist. The processes involved in the BCSP have been described in detail previously.

Specific data from the national database were retrieved on several KPIs: the caecal intubation rate (CIR), the adenoma detection rate (ADR), the mean number of adenomas detected per patient (MAP), the number of procedures performed by each colonoscopist and the comfort score for each procedure.

Data were also obtained from the accreditation database to provide data from the application forms submitted by colonoscopists detailing their performance in the 12 months prior to the application, including lifetime colonoscopic experience, the number of colonoscopies and polyp detection rate in the year prior to applying to become an accredited bowel cancer screening colonoscopist.

4.4.4 Statistical analysis

For analysis, colonoscopists were grouped into quartiles of performance for each of the KPIs separately. Each colonoscopist was therefore allocated 4 separate quartile performance

numbers for CIR, ADR, MAP and severe discomfort. Quartile 1 was nominated as the 'best-performing' quarter of screeners and therefore consisted of screening colonoscopists with the highest CIR/ADR/MAP and lowest proportion of patients with severe discomfort.

The DOPS data for all colonoscopists was analysed separately, with correlation between individual scores, overall scores and KPIs sought. The total DOPS scores for each individual domain across both assessors and both cases were added to give a maximum score of 16. The performance of colonoscopists with this maximum score was compared against the remainder of the cohort to assess whether a perfect score in a particular domain correlated with higher clinical performance.

A separate analysis examined data for candidates who failed their initial attempt at bowel cancer screening accreditation. Key performance indicators from these screeners were compared against those who passed on the first attempt.

The mark that candidates achieved to pass the MCQ was also analysed. Candidates were stratified into groups attaining over or under 80 per cent (80% being mid-way between 100% and the pass mark of 60%). KPIs were calculated for the two cohorts. The reported polyp detection rate from the logbook was correlated against the adenoma detection rate in the first 12 months of screening.

Data were analysed using SPSS version 20. The KPIs were individually tested with histogram plots and where these were equivocal the Kolmogorov – Smirnov test was calculated as a test of normality. The Kruskal-Wallis test was used to evaluate differences between multiple groups. The Mann-Whitney U test was used to compare independent groups with Spearman's rank

correlation coefficient calculated for correlation between variables that were not normally distributed, including CIR (figure 4.1). Pearson's correlation coefficient was calculated for correlations with the ADR (figure 4.1) and MAP, both of which were normally distributed. Spearman's correlation coefficient was used for calculations involving the CIR.

4.5 Results

4.5.1 Participants

The database was interrogated to obtain details of all 139,363 procedures performed between June 2006 and March 2012 by 244 screening colonoscopists. The data were filtered to include only procedures performed in the 12 months after each colonoscopist had commenced screening within the BCSP. 143 procedures recorded without a named endoscopist or by endoscopists who had not completed 12 months of screening activity were discarded, leaving 33,384 procedures in the evaluation group.

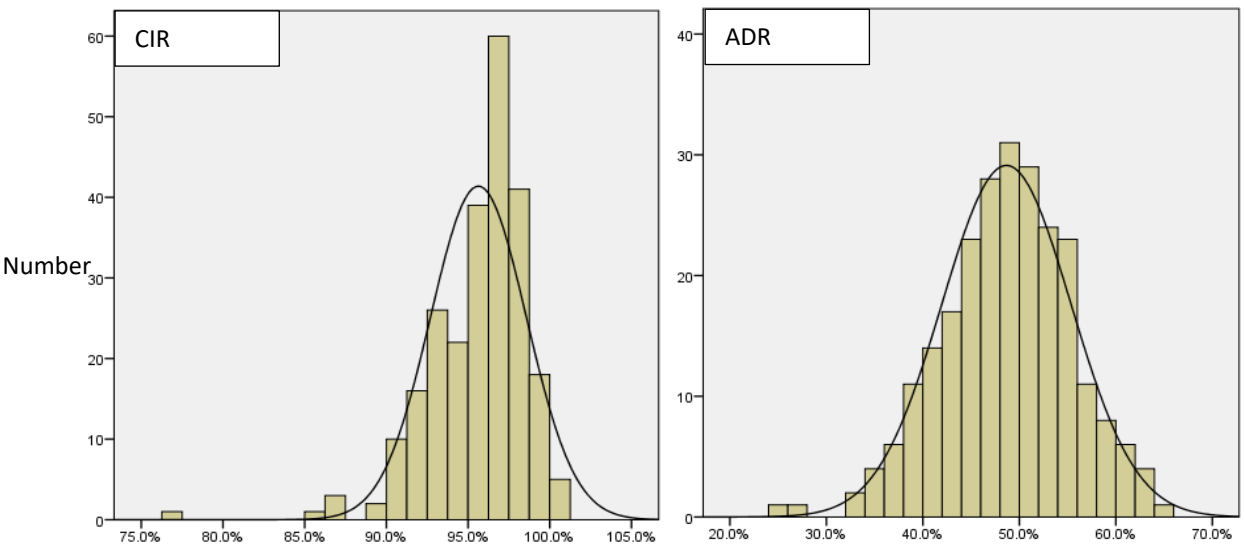

Figure 4.1. Caecal intubation rate and adenoma detection rate frequency histograms showing skewed and normal distributions respectively

4.5.2 Overall 12 month data

Table 4.1 shows baseline data amongst all 244 screeners, with minimum performance, maximum performance and mean performance in each quartile.

Table 4.1. Baseline data in first 12 months screening activity

	Max	Q1 mean	Q2 mean	Q3 mean	Q4 mean	Min	Overall mean
CIR / %	100	98.4	97.0	95.3	91.8	76.9	95.7
ADR / %	64.2	56.7	51.0	46.8	40.2	25.2	48.7
MAP	2.17	1.41	1.09	0.93	0.73	0.43	1.04
Severe discomfort / %	7.8	3.9	1.7	0.7	0	0	1.6

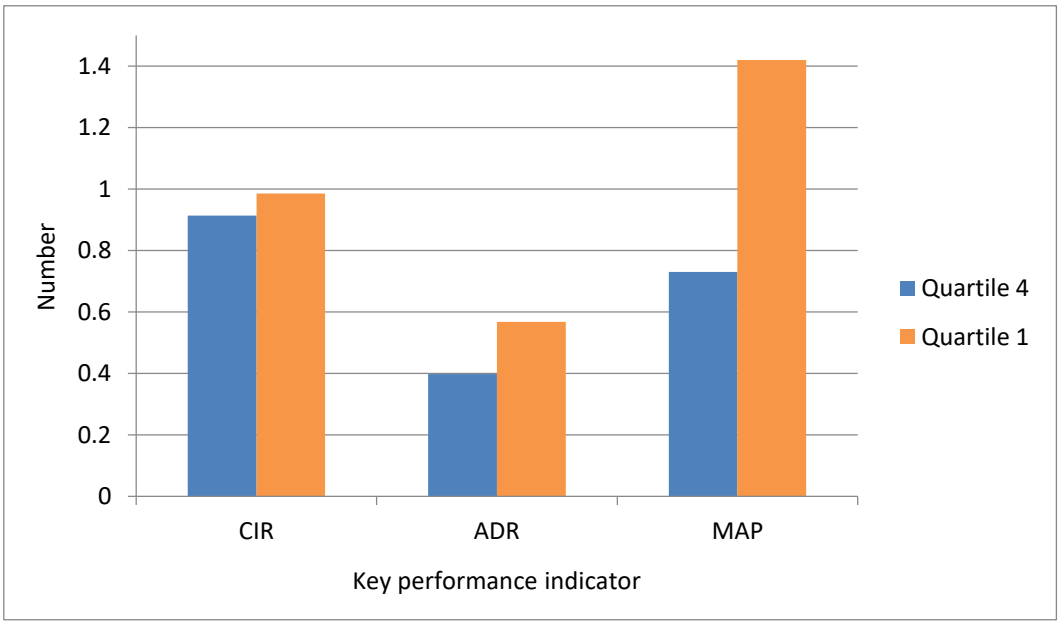

Figure 4.2. Performance of highest and lowest quartiles by KPI

4.5.2.1 Caecal intubation rate

Analysis of the CIR between the quartiles showed significant differences (Kruskal-Wallis test, p=0.003).The mean CIR in the top quartile was 98.4% compared to 91.8% (Table 4.2) in the lowest quartile (p<0.001).

There was a significant difference in the number of procedures performed in the year prior to commencing screening between the highest and lowest quartiles (271 compared to 216, p<0.01). The mean number of procedures in the first year of screening in the BCSP was 141 and 123 respectively (p=0.039).

Across all colonoscopists, there was a weak correlation between the number of screening procedures performed within the study period and CIR (Spearman's ρ=0.164, p=0.01). Data regarding the number of endoscopic procedures performed outside the BCSP during the period analysed were not available.

There was no significant difference between the CIR quartiles in the ADR, colonoscopy comfort, years of colonoscopy experience or total number of lifetime procedures performed. There was no correlation between CIR and ADR (Pearson's r=0.001, p=0.982).

Table 4.2. Analysis of colonoscopist performance stratified by caecal intubation rate

Criterion	Mean of top quartile for CIR	Mean of lowest quartile for CIR	Kruskal-Wallis test for differences between all quartiles	Mann Whitney U test between quartiles 1 and 4	Spearman's correlation coefficient p value (ρ)
ADR/%	49.7	48.6	0.490	N/A	0.180
MAP	1.11	1.30	0.098	N/A	0.074
Procedures performed in first 12 months of screening	141	123	**0.039***	**0.013***	**0.01*(0.164)**
Patients with severe pain/%	1.19	1.32	0.108	N/A	0.664
Procedures performed in year prior to commencing screening	271	216	**0.003***	**0.029***	**0.005*(0.185)**
Years of experience in colonoscopy	6.2	5.7	0.202	N/A	0.353
Number of lifetime procedures	2802	2707	0.565	N/A	0.268

4.5.2.2 Adenoma detection rate

In this separate analysis, colonoscopists were divided into quartiles by performance for ADR, with the highest performing colonoscopists in the first quartile and those with the lowest ADR in the fourth quartile. The mean ADR in the top quartile was 56.7% compared to 40.2% in the lowest quartile ($p<0.001$).

There was no significant difference between performance in all four ADR quartiles with respect to CIR, pain or any other measure from the lifetime procedure log. There was a significant difference in the MAP (1.64 compared to 0.77, $p<0.01$). Overall there was a strong correlation between ADR and MAP ($r=0.59$, $p<0.001$), although this was expected as ADR is linked to MAP.

Table 4.3. Analysis of colonoscopist performance stratified by ADR quartile

Criterion	Mean of top quartile for ADR	Mean of lowest quartile for ADR	Kruskal-Wallis test for differences between all quartiles	Mann Whitney U test between quartiles 1 and 4	Spearman's correlation coefficient p value (ρ)
CIR/%	95.5	94.8	0.234	N/A	0.18
MAP	1.64	0.77	<0.001*	<0.001*	<0.001 (0.815)
Procedures performed in first 12 months of screening	130	134	0.534	N/A	0.974
Patients with severe pain/%	1.48	1.61	0.884	N/A	0.852
Procedures performed in previous year	279	263	0.546	N/A	0.748
Years of experience in colonoscopy	7.1	7.3	0.781	N/A	0.855
Number of lifetime procedures	2677	3056	0.522	N/A	0.974

A funnel plot was also created using the adenoma detection rate data. This revealed that 7 endoscopists performed 3 standard deviations below the mean and 27 endoscopists performed at 2 standard deviations below the mean.

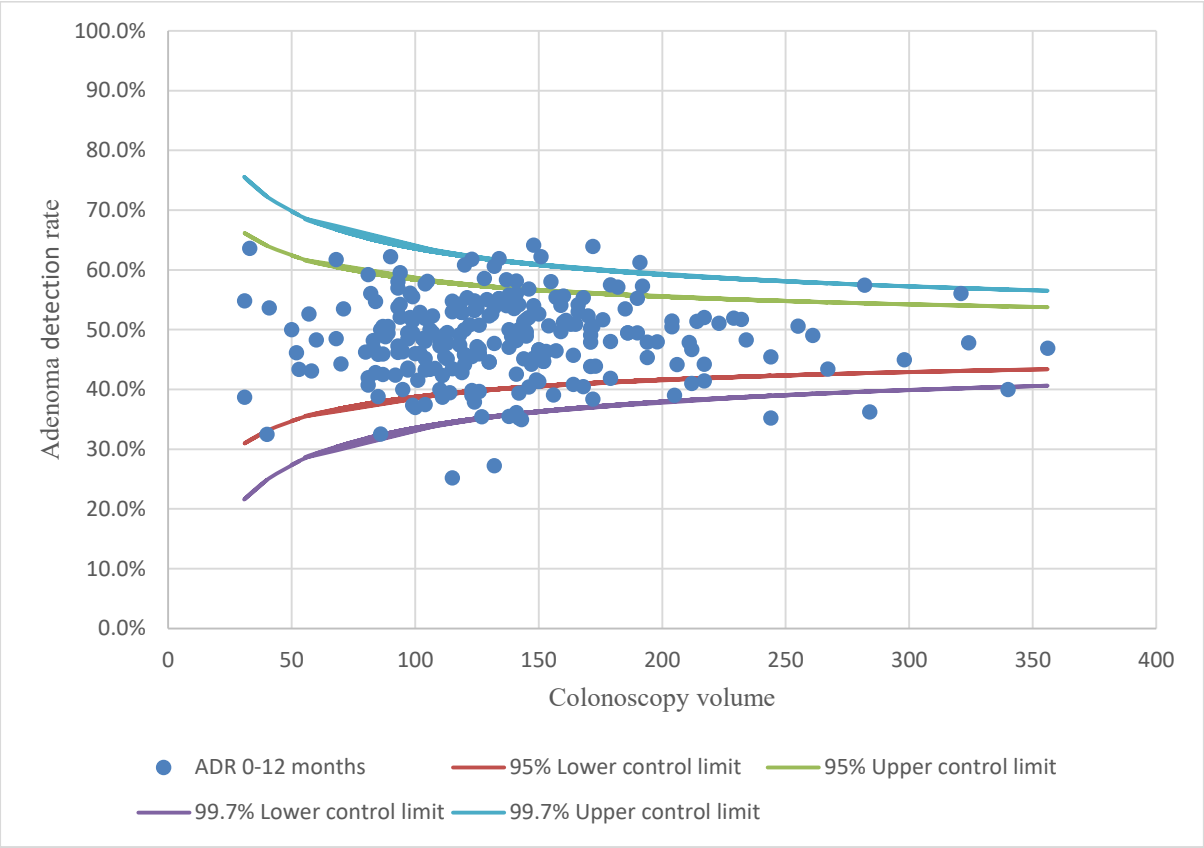

Figure 4.3. Funnel plot of adenoma detection rate for first 12 months screening activity

Analysis of the self-reported logbook polyp detection rate in the year before commencing screening showed a large range from 9% to 90%. There was a weak correlation between the polyp detection rate and the adenoma detection rate in the first 12 months of screening (Pearson's r=0.172, p=0.01, R^2 for trend=0.04).

4.5.2.3 Discomfort

Severe pain was an infrequent occurrence in many colonoscopists' cases. The best performing quartile of colonoscopists, categorised by the proportion of patients with severe pain, did not have any cases at all of severe pain during the 12 month study period.

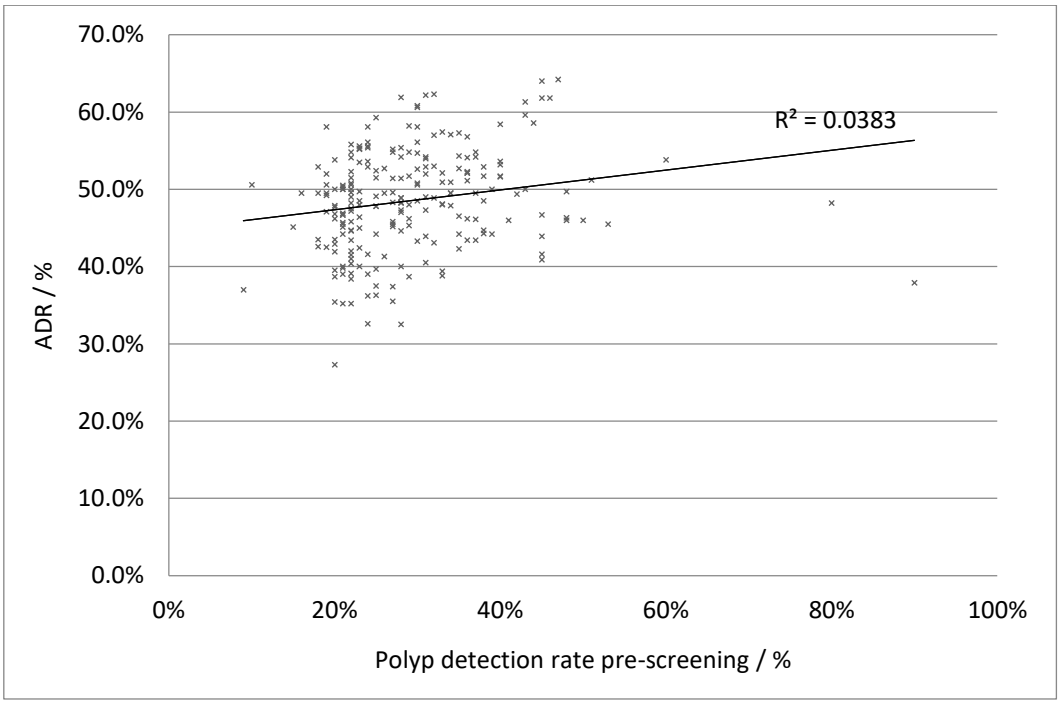

Figure 4.4. Polyp detection rate in the 12 months before screening compared to ADR in the first 12 months of screening

The poorest performing quartile, as defined by the quarter of colonoscopists causing the most discomfort, caused severe pain in a mean of 3.6% of patients. There was a significant difference between the quartiles (Kruskal-Wallis test, $p<0.001$).

There was a significant difference in ADR, CIR and MAP between the groups. This was most pronounced for the MAP where those causing least pain detected on average 1.44 polyps per patient against 0.99 in the group causing the greatest discomfort. The volume of procedures

performed in the past or during the year of screening evaluated did not influence the proportion of patients with severe pain.

Table 4.4. Analysis of colonoscopist performance stratified by proportion of patients with severe pain

Criterion	Mean of top quartile for severe pain (best performers)	Mean of lowest quartile for severe pain (worst performers)	Kruskal-Wallis test for differences between all quartiles	Mann Whitney U test between quartiles 1 and 4	Spearman's correlation coefficient p value (ρ)
CIR/%	96.3	95.6	**0.027***	0.083	0.664
ADR/%	50.5	48.8	**0.027***	0.197	0.852
MAP/ per patient	1.44	0.99	**0.033***	**0.012***	**0.05 (-0.125)**
Procedures performed in first 12 months of screening	128	129	0.066	N/A	0.886
Procedures performed in previous year	271	259	0.547	N/A	0.616
Years of experience in colonoscopy	6.0	6.7	0.325	N/A	0.168
Number of lifetime procedures	3014	2498	0.938	N/A	0.893

4.5.3 Accreditation data

230 colonoscopists (94%) passed the accreditation on the first attempt; 11 required two attempts and three candidates passed on the third attempt. No data were available for those endoscopists who attempted the accreditation but never succeeded and hence never

commenced screening. Colonoscopists who were successful on the first attempt were compared against the 14 others.

4.5.3.1 Passing second/ third time

Data for the 14 first unsuccessful attempts at accreditation, equating to 28 colonoscopies (two per candidate) and 56 DOPS assessments (two per colonoscopy), were analysed. The worst scoring domain was for loop recognition and resolution: 51.8% of assessments were marked inadequate. The second worst domain pertained to the use of torque steering, with 46.4% assessments scoring one or two on the DOPS marking scheme.

There was no significant difference in CIR, ADR, MAP or comfort in the first 12 months between those who passed first time and those who required more than one attempt at accreditation before passing (Table 4.5).

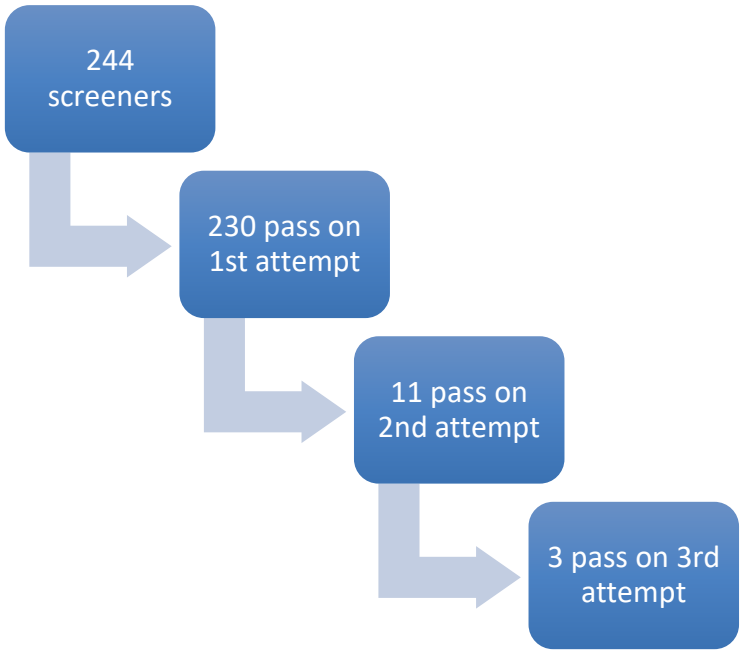

Figure 4.5. Number of screeners passing on each attempt at accreditation

Table 4.5. Analysis of KPIs for colonoscopists who passed accreditation on the first opportunity versus others

	Pass on first attempt	Did not pass on first attempt	p value
CIR / %	95.5	96.5	0.14
ADR / %	49.0	47.4	0.62
MAP	1.13	0.92	0.56
Severe pain / %	1.4	1.9	0.35

4.5.3.2 No accreditation data cohort

A total of 18 screening colonoscopists did not sit the formal accreditation process. These endoscopists were part of the inception cohort of screeners, recruited before the accreditation process had been devised. Performance data from these individuals were compared against the other 226 colonoscopists. Those that took the accreditation had significantly higher caecal intubation rate of 95.8% versus 94.0% (Mann-Whitney U test, p=0.01). There were no other differences in the rates of polyp detection or of severe discomfort caused.

Table 4.6. Analysis of KPIs for colonoscopists who started screening before the formal accreditation process was introduced

	Did not take accreditation	Took accreditation	p value
n	18	226	
CIR / %	94.0	95.8	**0.01***
ADR / %	49.4	48.5	0.59
MAP	1.15	1.03	0.30
Severe pain / %	1.4	1.4	0.91

4.5.4 DOPS analysis

4.5.4.1 Mucosal visualisation and ADR

48 candidates scored perfectly in the "adequacy of mucosal visualisation" domain. These candidates were compared against the other 178 colonoscopists. There was a non-significant trend towards a better ADR (49.8 per cent versus 48.2 per cent, p=0.059).

A separate domain examines maintenance of the luminal view on extubation. The candidates who maintained a perfect luminal view throughout the assessment did not perform better with respect to ADR or MAP (t test, p=0.058 and 0.070 respectively). There was no difference in those who had been judged to use position change perfectly on extubation with respect to ADR or MAP (t test, p=0.355 and 0.823 respectively).

4.5.4.2 Comfort

The 126 colonoscopists that scored perfectly for "treating the patient with respect" had significantly more comfortable patients. No, minimal or mild pain was experienced by 88.8% of patients in this cohort versus 86.8% of patients in the other group (Mann Whitney U test, p=0.032). The domain examining candidates' awareness of patients' pain was tested to ascertain whether higher performance in the accreditation predicted comfort during screening practice. There was also no difference between those scoring perfectly with severe pain (Mann Whitney U test p=0.997 respectively).

There was no difference in the comfort of patients after analysis of domains pertaining to distension (p=0.117), torque (p=0.086), loop recognition (p=0.084) or the administration of sedative (p=0.681).

4.5.5 MCQ analysis

The group scoring over 80 per cent included 130 colonoscopists. These individuals performed significantly better than the others with respect to ADR, MAP and severe pain (Table 4.7).

Table 4.7. Correlation of high-scoring colonoscopists on MCQ against KPIs

MCQ score	Under 80%	Over 80%	p value
CIR / %	95.7	95.7	0.997
ADR / %	47.6	49.3	0.048*
MAP	0.98	1.08	0.004*
Severe discomfort / %	1.6	1.3	0.044*
Years in post	6.9	6.0	0.739
Lifetime colonoscopy experience	2777	2477	0.936

4.6 Discussion

4.6.1 Do factors from accreditation predict performance in first 12 months?

This study was unable to accurately predict how individually screeners would perform using data from the initial accreditation. Those who were not successful during the first attempt at accreditation managed to perform similarly to other colonoscopists with respect to rates of procedure completion (CIR), adenoma detection (ADR/MAP) and discomfort, which lends credence to the validity of the accreditation process.

The corollary of this however is that no data were available on those who attempted the screening accreditation process and failed yet either did not re-attempt the test or failed on the second attempt. It would be a logical inference to hypothesise that the performance of

these individuals would not be as good as those who had passed the stringent accreditation, although this supposition is not backed by data.

Interestingly, when data from the inception cohort of screeners, largely including those considered experts, were analysed (Table 4.6**)**, there was a statistically poorer performance seen with regards to procedure completion in this group but KPIs pertaining to adenoma detection and comfort were not different from those who had undertaken the formal accreditation. Given these experts have largely performed as well as later screeners, this does raise questions about the necessity of the accreditation process and whether it should apply to all endoscopists. Chapter 6 explores some of the characteristics of expert endoscopists; it may be in the future some of the attributes identified could be used to target an accreditation at specific individuals.

Individual parameters from the DOPS did not correlate strongly with or predict future key performance indicators. This is likely to be as the DOPS assessment tool used to mark assessments is not sensitive enough to differentiate exceptional from competent performance; indeed it was devised for the documentation of technical competence at colonoscopy.

A previous study has looked specifically at the performance of the whole accreditation process using G-theory: this work found similar reasons for colonoscopists failing the accreditation and showed a high level of robustness when the accreditation process considered in its entirety[70].

The accreditation DOPS most rigorously assesses technical skills involved in insertion, although the entire colonoscopy procedure is evaluated. There was only a weak correlation

between adenoma detection rates in the first year of screening activity compared to polyp detection rates in the 12 months prior to the commencement of screening. Data from the logbook are likely to be less reliable, as evidenced by the very large range in performance documented, than the independently recorded data from the BCSP itself and the poor correlation may be a reflection of this.

The strongest predictor of performance came surprisingly from the MCQ test, whereby those who scored over 80% - significantly in excess of the pass mark of 60% - had higher rates of adenoma detection and more comfortable patients. The reasons for this association are unclear as analysis of the number of lifetime procedures performed and number of years as a fully qualified consultant showed no difference between those just passing and those scoring very highly.

There are some data that the declarative knowledge used in tests such as MCQs can be subconsciously linked to the complex procedural knowledge required to efficiently perform a task such as colonoscopy[104,105]. In other domains such as mathematics, conceptual knowledge has been shown to be convincingly linked to the ability to utilise the correct procedure for solving a mathematical problem[106]. No literature links non-technical skills with declarative knowledge but this avenue has remained largely unexplored to date.

4.6.2 How does performance vary in the first 12 months screening?

The vast majority of screening colonoscopists perform well above the rigorous minimum standards set in the BCSP.[107] However, even in a selected group of high-performing

colonoscopists with significant experience of colonoscopy, there is considerable variation in performance in CIR, ADR, MAP and patient comfort.

There was no correlation between many key performance indicators used to monitor endoscopists, including CIR and ADR. There was however a significant association between CIR and the endoscopic activity the previous year, suggesting a benefit of higher volume on performance. Colonoscopists performing more procedures in their first year screening patients did not however have higher adenoma detection rates. These data lend weight to the practice of collecting data spanning a large number of performance metrics as this study shows it is not possible to extrapolate performance in one domain from that in another.

These data did demonstrate an association between screening case volume and caecal intubation rate. However, the reliability of this association is unlikely to be robust as the difference in the number of cases between the highest and lowest performing groups was modest (18 cases over a year) and the analysis did not take into account the number of additional colonoscopic procedures performed outside the BCSP during the study period.

Overall, there was no association between case volume within the screening programme and ADR or MAP. A recent study looking at the factors influencing the quality of screening colonoscopy and adenoma detection rates in Germany found similar results[108]. These findings differ from those encountered in surgery, where higher volumes have been linked to significantly better outcomes both individually for surgeons and in combination for hospitals[109] but may add weight to the recommendation to perform a minimum number of procedures per annum.

In this cohort, colonoscopists most likely to cause severe discomfort during the procedure had worse caecal intubation rates and adenoma detection rates than their peers performing better tolerated examinations. This points to technical ability playing an important role in achieving a high quality examination with minimal discomfort; however, some potentially confounding factors such as geographical variation and the ethnic origin of patients were not taken into account.

4.6.3 Strengths / limitations

This study is the largest in the literature examining real-world performance characteristics of a very large cohort of accredited colonoscopists performing endoscopy on a homogenous group of faecal occult blood positive patients within a narrow age band. There was no selection bias as all the colonoscopists practising under the BCSP were included in the analyses. The study adds some weight to the validity of the entire accreditation process as most but not all colonoscopists performed well against target key performance indicators. Differences in performance may be expected given the size of the cohort of screening colonoscopists. It is unclear whether these differences have any clinical implications.

This study was a retrospective data analysis and specifically did not examine patient or technical factors related to colonoscopy, such as variability in endoscope performance. Previous studies have shown that sex, age, social deprivation and smoking all have an effect on the development of colonic polyps[110]. Previous work has shown that these factors do not introduce significant bias in the assessment of colonoscopist performance with the BCSP[107]. There is ongoing debate in the literature about the true effect of colonoscopic withdrawal

time with trials reporting conflicting results. Trials examining the use of hyoscine butylbromide on adenoma detection rate have also yielded inconclusive results[111,112].

Most of the outcome measures used in this study, such as caecal intubation rate and adenoma detection rate, are objective. However, the assessment of pain is a subjective process and has not been validated and there is consequently likely to be variability in the scoring of this between specialist practitioners. This may also be compounded in centres where the same specialist screening practitioner enters the data for all procedures performed by individual endoscopists where bias is more likely. Previous studies have shown significant differences between reported comfort scores between endoscopists, nurses and patients.[113]

4.6.4 Summary

In summary, this large study could not identify strong predictors of colonoscopic performance from a multi-stage accreditation test. Performance was generally very high across all individuals, but a discernible performance gap exists, the cause of which is not clear. Further work will try to identify potential causes of these differences in performance.

5 Screening colonoscopist performance variation over time in the English Bowel Cancer Screening Programme

5.1 Background

Chapter 4 has described how performance varied in colonoscopists new to the BCSP, with results in keeping with other studies of cross-sectional performance amongst colonoscopists[114]. However, the effects of the passage of time on individual colonoscopy performance have not been assessed to date in the literature.

A variety of studies and guidelines have been published (section 1.5.1.1) detailing a consensus in the literature as to the approximate number of procedures required to develop de novo competency, the volume required to maintain performance in those undertaking screening colonoscopy is unclear.

5.1.1 Importance of longitudinal data

Longitudinal performance at colonoscopy is of particular interest in colorectal cancer screening as the importance of key performance indicators such as adenoma detection rate have been shown to be linked to interval cancer rates[1]. Identification of those performing less well than their colleagues could be used to target interventions to improve quality across a screening programme.

Previous work examining adenoma detection rates at screening flexible sigmoidoscopy has shown that lower detectors' performance persisted despite gaining experience and receiving regular feedback[115]. It is unknown whether similar effects are seen with screening colonoscopy.

5.1.2 Expertise

Colonoscopists performing more strongly initially in a selected group could reasonably be regarded as experts. Whether enhanced performance metrics are maintained in this group from year to year has never been explored and raises the question about whether there is something innate about these individuals that leads to exceptional performance or whether this can be attributed to random chance.

5.2 Aims

5.2.1 Primary outcome

5.2.1.1 Effect of time on performance

The principal aim was to assess how performance varied over time in this highly selected group of accredited colonoscopists in the years after they commenced screening in order to ascertain whether there was any evidence of performance fatigue, with performance falling over time or the converse with performance improving with volume over time.

5.2.2 Secondary outcomes

5.2.2.1 Threshold volume

The secondary aim was to determine whether there was a threshold number of procedures to be performed per annum within the screening programme that was needed to maintain performance across important KPIs.

5.3 Hypothesis

It is predicted that the best performers will remain so throughout the study period. It is also hypothesized that those endoscopists performing the greatest number of procedures will perform more strongly.

5.4 Methods

5.4.1 Approvals

Endoscopists in the Bowel Cancer Screening Programme give their consent to their anonymised data being analysed for research purposes. The formal study proposal was peer reviewed by the Bowel Cancer Screening Programme Research Committee, who gave formal permission for the study.

5.4.2 Participants

Data were collected from June 2006 to March 2012 on screening colonoscopists who had completed 36 months of screening activity, with a minimum activity level of 30 procedures per annum for each 12 month period. Procedures without a named colonoscopist were discarded.

5.4.3 Data collection

The same central BCSP database was utilised to collect data on several KPIs: caecal intubation rate (CIR), the adenoma detection rate, the mean number of adenomas detected per patient (MAP), negative colonoscopy withdrawal time (NCWT), patient comfort scores and annual procedure volume. The negative colonoscopy withdrawal time was defined as the time taken

to withdraw the colonoscope from the patient, starting at the caecum, for procedures where no polyps were detected.

Data were also collected on the year colonoscopists started participating within the screening programme.

5.4.4 Statistical analysis

Colonoscopists were stratified into quartiles for CIR, ADR and MAP by their performance within the first year of screening. Each colonoscopist was therefore assigned three separate quartile numbers for each of these three KPIs. The performance of each quartile of colonoscopists was then analysed over the three year study period for each KPI, based on their performance within their initial 12 months of screening activity.

For CIR, an additional analysis was performed. Colonoscopists were divided into two groups: the first group comprised those that did not attain the BCSP quality benchmark specifying a CIR of greater than 90 per cent in the first 12 months; the second group comprised all other colonoscopists. The performance of the two groups was then compared over time.

To assess whether the date at which colonoscopists started screening influenced performance, they were separately divided into groups by the year that they first performed a screening procedure. KPIs for each of these year groups were analysed to determine whether the newer entrants to the programme affected the results.

For analysis of the ADR, the NCWT was also compared across quartiles to assess whether this was responsible for differences in detection rates

The distribution of performance across the three KPIs previously tested was used to determine the statistical test. The Mann-Whitney U test was used to compare non-parametrically distributed KPIs and the t test used to compare normally distributed data. The Kruskal-Wallis test was used to compare data across quartiles.

5.5 Results

5.5.1 Participants and procedures

There were 139,363 procedures performed between June 2006 and March 2012 by 154 colonoscopists. The 391 procedures logged without a named colonoscopist or by colonoscopists who had performed fewer than 30 procedures per annum under the programme were excluded.

5.5.2 Caecal intubation rate

5.5.2.1 Overall

In screeners' first year, the mean CIR across all colonoscopists was 95.3 per cent (range 76.9%-100%). In the third year, this had improved to 96.8 per cent (Figure 5.1, $p<0.001$).

5.5.2.2 Analysis of those with CIR less than 90 per cent in year 1

The mean CIR for the seven endoscopists who did not meet the KPI standard of a 90 per cent CIR in the first year of screening was 86.1 per cent; all remained in the programme and by the third year, this had improved to exactly match the performance of the rest of the group, attaining a CIR of 96.8 per cent. There was no statistical difference at three years between the two groups (Mann Whitney U-test, $p=0.93$).

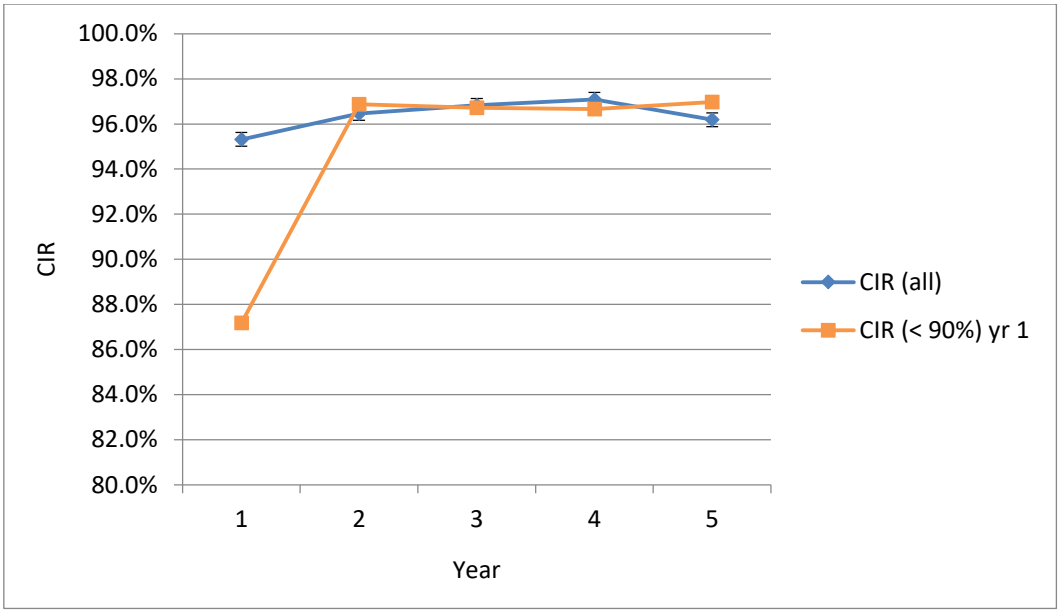

Figure 5.1. CIR showing performance of those with a CIR of under 90 per cent in year 1 against others

5.5.2.3 Quartile analysis

When the entire cohort was divided into quartiles of performance based on the CIR in the first 12 months of activity, the highest performing quartile (Q1) had a mean CIR of 98.4% compared to Q4 which had a mean CIR of 90.9%.

At three years, the Q1 CIR was 97.6% compared to 95.8% for Q4 ($p<0.001$). There was a small fall of 0.8% in the CIR of the top quartile ($p<0.001$) and a rise of 4.9% in the lowest quartile ($p<0.001$). The quartiles maintained their ranking across the three year study period.

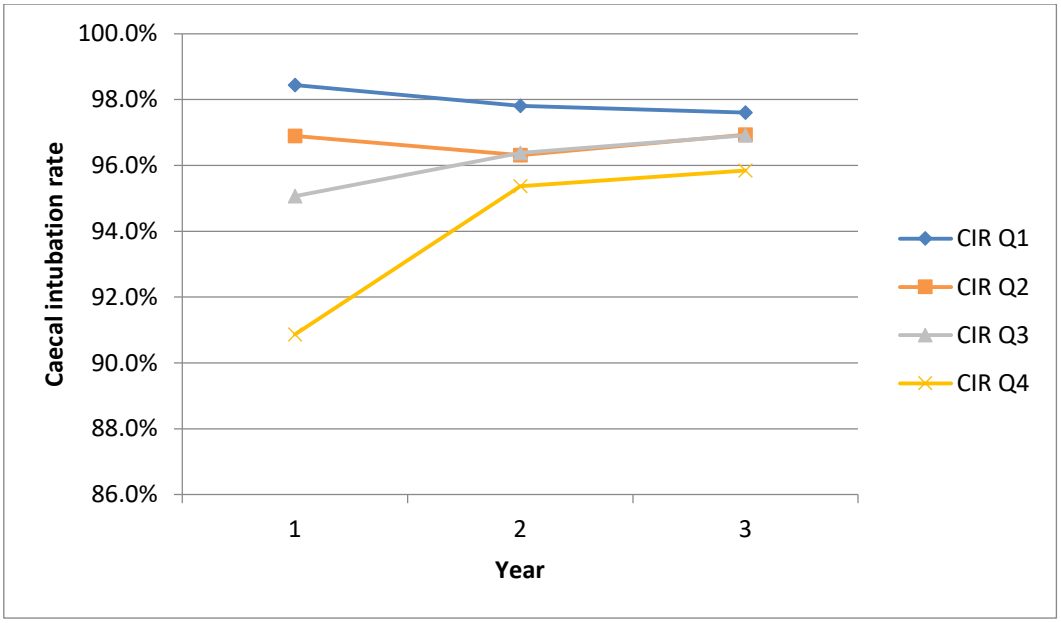

Figure 5.2. CIR performance against time categorised by quartiles

5.5.3 Adenoma detection rate

5.5.3.1 Overall

The overall ADR amongst all colonoscopists fell from 48.8 per cent to 44.6 per cent ($p<0.001$) from year one to year three.

5.5.3.2 Quartile analysis

Colonoscopists in the lowest quartile (Q4) with respect to ADR were compared to the highest quartile (Q1). In the colonoscopists' first year, the mean ADR was 40.2 per cent in the lowest quartile and 56.8 per cent in the highest quartile. At three years, the ADR was 39.9 per cent compared to 49.5 per cent ($p<0.001$).

There was no significant change in the ADR for the lowest quartile over this period (p=0.689) whereas the fall in ADR for the top quartile was significant (p<0.001). Again, the quartiles maintained their ranking throughout the three years of the study (Figure 5.3).

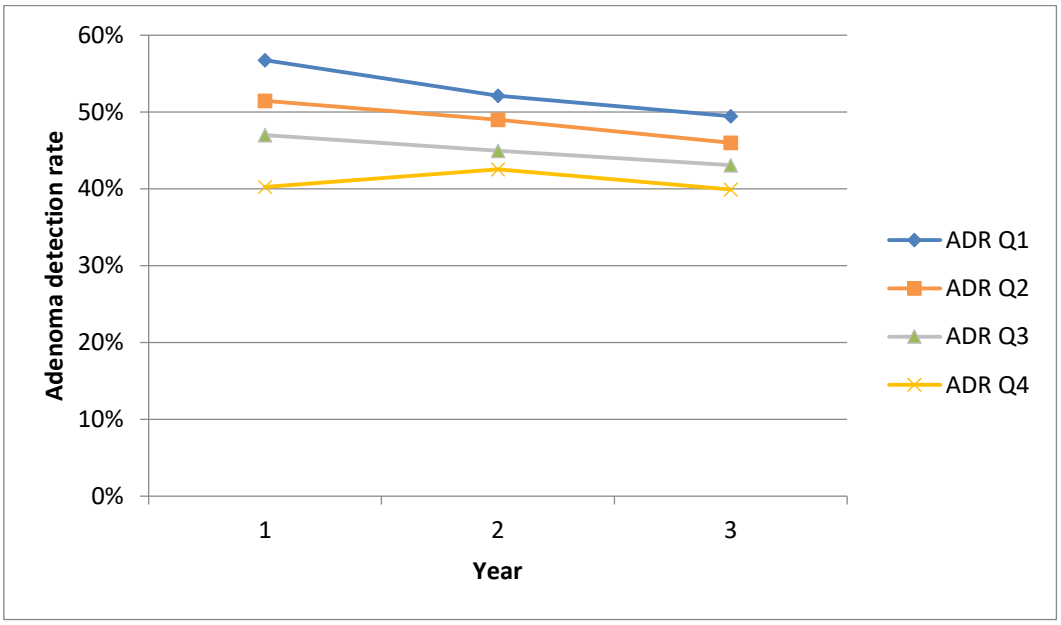

Figure 5.3. ADR against time for each quartile

5.5.3.3 Withdrawal time

There was no difference between the NCWT between the first quartile and the fourth quartile in any of the three colonoscopist years of the study period (Table 5.1).

Table 5.1. Negative colonoscopy withdrawal times for colonoscopists with highest and lowest ADR

Year	ADR Quartile	Mean NCWT / minutes	p value
1	1	10.6	0.623
	4	9.6	
2	1	9.6	0.782
	4	9.5	
3	1	9.4	0.162
	4	8.6	

5.5.4 Mean number of adenomas per patient

5.5.4.1 Overall

In the first year, the MAP was 1.04 (range 0.43-2.17). By year three, this had fallen to 0.98 (range 0.46-2.10). The fall in the number of adenomas detected overall was significant (paired t test, p=0.011).

5.5.4.2 Quartile analysis

Colonoscopists in the lowest quartile with respect to MAP were compared to the highest quartile. The MAP in the highest quartile fell from 1.41 to 1.30 (paired t test, p=0.043), whereas in the lowest quartile the MAP rose from 0.72 to 0.79 (paired t test, p=0.047). Again, the colonoscopist quartiles remained in the same order throughout the study period.

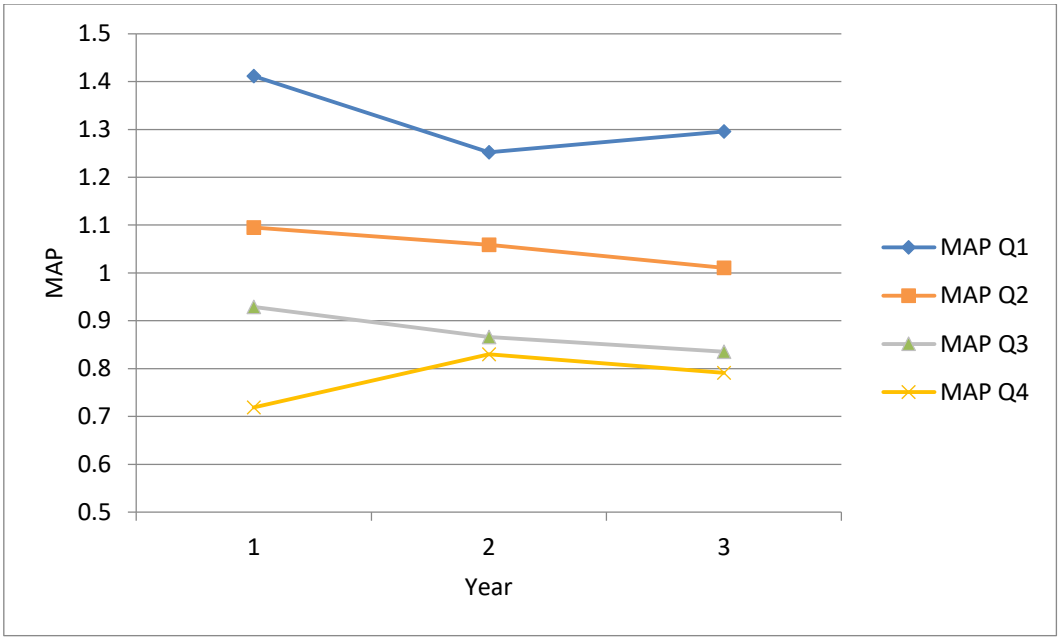

Figure 5.4. MAP against time for each quartile

5.5.5 Comfort

5.5.5.1 Overall

Overall, the proportion of patients with none, minimal or mild discomfort was 87.7 per cent in screeners' first year of practice, rising to 90.5 per cent in year three ($p<0.001$). Similarly, the proportion of patients with severe discomfort fell from 1.4 per cent (range 0-6.0) to 1.1 per cent ($p=0.024$).

5.5.5.2 Quartile analysis

Data pertaining to the quartile of colonoscopists causing the most severe discomfort in the first year (affecting a mean of 3.9 per cent of patients) were compared against the best performing quartile, causing least discomfort (no patients suffering from severe discomfort for any of the colonoscopists' procedures). After three years, 2.3 per cent of patients had

severe discomfort in the highest quartile, compared to 1.2 per cent in the lowest quartile (p=0.001). In the third year, quartile 1 caused a higher level of pain than quartile 2, but this difference was not statistically significant (p=0.228).

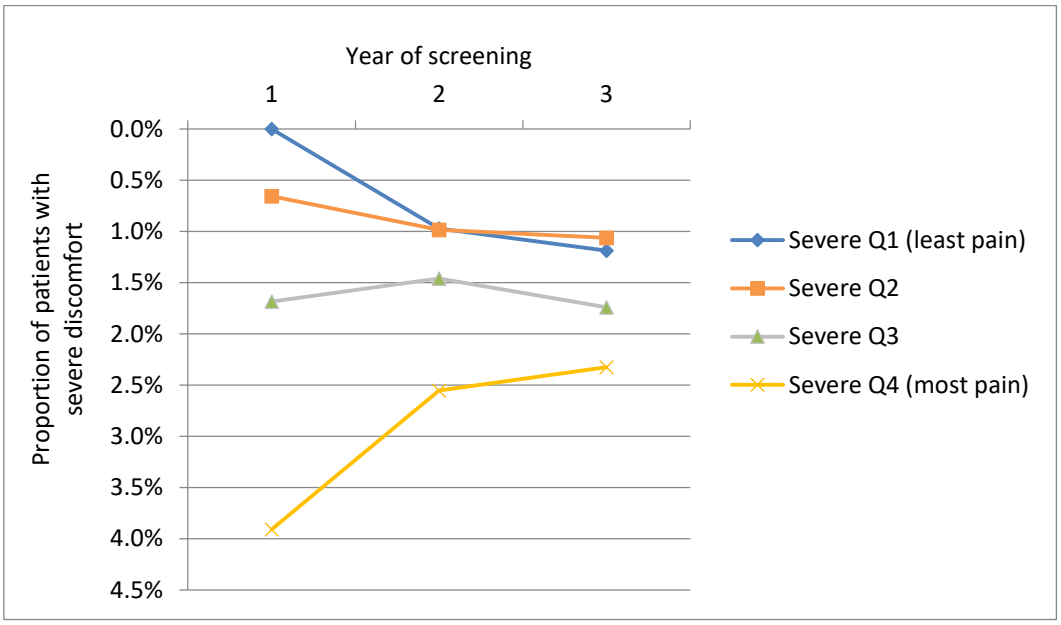

Figure 5.5. Severe discomfort against time for each quartile

5.5.6 Volume

5.5.6.1 Overall

The mean number of procedures performed in the first year was 136 (range 31-324), rising to 146 by year three (range 39-434).The majority of colonoscopists (70.4 per cent) did not attain the target number of screening procedures.

5.5.6.2 Volume against performance

To assess the impact of volume and performance, colonoscopists were initially categorised into two groups: those who had performed fewer than 150 procedures in their first 12 months

and the remainder of the cohort. There was no significant difference between the two groups with respect to the CIR (96.2% vs 95.3%, p=0.181), ADR (48.9% vs 48.4%, p=0.361), MAP (1.04 vs 1.03, p=0.285) or severe pain (1.2% vs 1.6%, p=0.736) in the first 12 months.

A further analysis examined colonoscopists who performed fewer than 100 procedures in each of the three years. These 11 colonoscopists were compared against 141 colonoscopists who had performed at least 100 procedures each year. There was no significant difference in CIR (93.1% vs 95.5%, p=0.32, 95.8% vs 96.5%, p=0,65, 96.8% vs 96.8%, p=0.99), ADR (48.5% vs 48.8%, p=0.90, 45.9% vs 47.2%, p=0.68, 45.0% vs 44.5%, p=0.60) or MAP (1.18 vs 1.03, p=0.51,0.98 vs 1.00, p=0.79, 1.15 vs 0.97, p=0.99) for any 12 month time period during the three years.

When the 24 colonoscopists who performed the recommended volume (150 procedures per annum) for each of the 3 years were compared against the 11 colonoscopists who had performed under 100 procedures per year for three years, there was no significant difference in any of the three years for CIR (p=0.25,0.39,0.88), ADR (p=0.74,0.17,0.75) or, MAP (p=0.34, 0.44, 0.15).

5.5.7 Calendar year of commencing screening

When screeners were divided into cohorts by the calendar year of commencement screening (2006-2009), there was a difference between the groups in respect of CIR (p=0.001; Table 5.2).

There was an increase in CIR in screeners commencing in 2008. Performance in this cohort was significant better than those starting in 2006 and 2007 (Mann Whitney U test, p=0.02 and

<0.0001 respectively). There was no difference in CIR between those commencing in the years 2008 and 2009 (Mann Whitney U test, p=0.1).

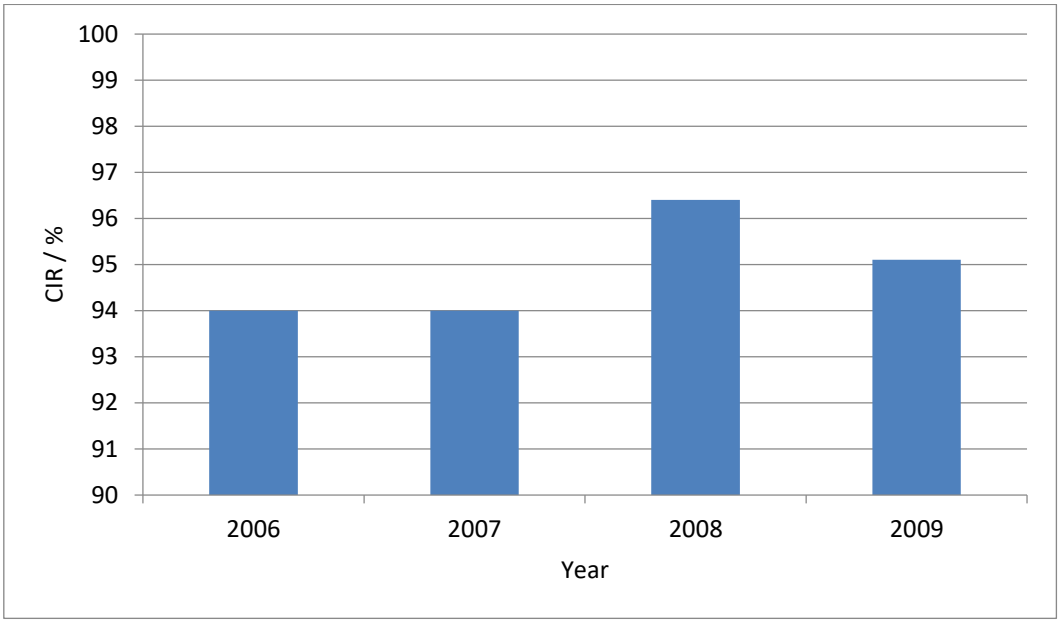

Figure 5.6. Mean CIR in first 12 months by year commencing screening

The majority of KPIs however showed no difference in performance for screeners joining the programme at a later stage. ADR, MAP, and comfort all did not change significantly over this period (Table 5.2).

Table 5.2. KPIs by calendar year commencing screening

	No of screeners	**CIR/%**	**ADR/%**	**MAP**	**Severe pain/%**
2006	14	94.0	48.4	1.19	1.2
2007	41	94.0	48.0	1.04	1.4
2008	73	96.4	49.0	1.01	1.3
2009	18	95.1	48.5	1.02	1.3
Total/mean	**146**	**95.3**	**48.8**	**1.04**	**1.4**
Kruskal-Wallis test p value		<0.001*	0.982	0.802	0.813

5.6 Discussion

These data reinforce the findings from Chapter 4 by showing that even in a group of accredited screening colonoscopists performance varies significantly within the group, over time and between different key performance indicators. Importantly, these differences are sustained, with the colonoscopists demonstrating the most expertise initially remaining the best performers after several years and with the lowest performers also retaining their ranking over time.

When the entire cohort of 154 colonoscopists were ranked according to their first year's caecal intubation rate performance, the performance quartiles maintained their ranking over time, despite a degree of "regression to the mean" with an improvement of 4.9% in CIR in the poorest performing quartile. However, the absolute difference in CIR after 3 years of screening was small (1.8%).

The improvement in the attainment seen for the poorest performers in CIR was not seen for KPIs related to adenoma detection. With respect to ADR, against a generally falling detection rate, those performing least well maintained their adenoma detection rates but still did not reach the same levels as those performing better initially. The general fall in ADR over time was an expected finding due to a higher incidence of neoplasia and polyps in the prevalent round of screening and of the burden of polyps in the screened population diminishing due to the screening programme.[116] Similar findings were found for MAP, with the lowest quartile remaining so throughout the period of analysis.

An alternative explanation for the falling adenoma detection rate might be fatigue from performing colonoscopy repetitively. The maintenance of a long negative colonoscopy withdrawal time suggests that colonoscopists were not accelerating during procedures due to tedium from performing colonoscopy over several years. However, it is not possible to conclusively exclude fatigue as an explanation for this phenomenon as withdrawal time is not the only variable associated with ADR[108]. These findings warrant thorough further scrutiny to conclusively exclude fatigue, as if this were the case, this would have profound implications with regards to service configuration and delivery.

With regards to discomfort, the vast majority of patients underwent a comfortable procedure with on average only between 1.1% and 1.4% experiencing severe discomfort. Those causing most discomfort improved over time but still maintained a significant divide after several years, never attaining the same performance levels as those causing least discomfort initially though the difference in absolute terms between the groups was small.

Why do these significant differences between performance quartiles occur and persist over time? Given that the cohort of FOBT positive patients is homogeneous, the most likely explanation is that of differing technical abilities amongst screeners.

Many publications have sought to examine the impact of withdrawal time from the caecum but different centres have yielded conflicting results[37,117,118]. A recent study which showed no impact on withdrawal time on ADR did however find significant differences between high and low adenoma detectors in endoscopic technique with regards to inspection behind folds, bowel distension and cleansing manoeuvres[119]. Other studies have shown the positive effect on ADR of position change during withdrawal[120].

This large study shows that although there is capacity for improvement in some quality domains of colonoscopy such as CIR, colonoscopists largely remain ranked in the same performance order regardless of the KPI measured. This has important implications for quality assurance and training. As the discrete performance hierarchy is stable over time, early intervention seems appropriate.

Given the paramount importance of ADR in improving the efficacy of screening colonoscopy[36] and the wide variation seen in this cohort of highly selected colonoscopists, it may be prudent to target additional training resources at those performing either below the mean ADR/MAP in an effort to homogenise the performance of the entire group or a standard deviation below the mean using a funnel plot (see Figure 4.3).

What remains to be seen is whether such training interventions will successfully bridge the performance hierarchy. Studies assessing interventions to improve ADR have largely been unsuccessful[121], suggesting either inadequate intervention or that the quality difference may be "hard-wired" and unbridgeable. Further work is needed in this area.

The analyses relating to volume showed no difference in the KPIs when comparing those who performed higher numbers of screening colonoscopies and those performing far fewer procedures. One of the limitations of this conclusion is the lack of inclusion of non-screening procedures performed by each individual. BCSP colonoscopists are likely to be regarded as specialist endoscopists within their respective units and perform regular colonoscopy in the symptomatic population in addition to their screening commitments. Although this is likely to vary considerably between individuals, it is likely that the combination of screening and symptomatic service requirements mean that all colonoscopists perform in aggregate

hundreds of procedures per annum. Variations in procedure volume at this level are unlikely to result in clinical differences in performance.

Although a specific volume may not be required to maintain performance on these criteria, other factors not analysed in this study, such as reliability of data, concentration of expertise, complication data, interval cancer, and completeness of polyp excision could conceivably be affected by procedure volume, but evidence for volume affecting these other outcomes is lacking.

Analysis of performance according to when colonoscopists undertook their first screening procedure showed a small but significant improvement in CIR in those commencing in 2008/9 compared to 2006/7. One explanation may be better recent training in England due to the introduction and uptake of certification of colonoscopy competency for all endoscopy trainees, as outlined in Chapter 2. Given the absolute difference in CIR was small, this could also be a chance observation unlikely to be of clinical significance.

The accreditation process for these later-starting screeners was identical to those who commenced earlier and so it is unlikely that changes in the entry criteria selected better performing individuals, although it remains possible that a desire to support prompt roll-out of the screening programme initially resulted in a subconscious lowering of accreditation threshold. It is interesting to note that this effect was only seen with CIR and not with other KPIs.

5.7 Conclusion

This study has shown that even in a high quality colonoscopy service, a discrete performance hierarchy exists, which remains unchanged over time for all key performance indicators. With the exception of ADR, which decreased as the programme moved from prevalent to incident rounds, overall performance improved over time, suggesting that performance fatigue is not a significant issue in the screening programme. It must be acknowledged that the absolute differences in performance between the cohorts, particularly with respect to caecal intubation rate, were small and although these differences meet statistical significance, this may be a function of the large size of the dataset. The clinical implications of these small differences has not been proven in trials.

These data do suggest that the best predictor of future colonoscopic performance in experienced endoscopists is current performance, yet what drives the very best colonoscopists to attain consistently high KPIs remains to be answered. It is possible to speculate that the personality of endoscopists may be partly responsible for performance differences, but there has been no work to date looking at individual personality or psychological attributes and linking them to endoscopic performance outcomes.

6 Factors defining expertise in screening colonoscopy

6.1 Background

Bowel cancer screening (BCS) has been successfully rolled out across England, with stringent quality requirements for units undertaking screening and individuals performing colonoscopy within the programme. There is a rigorous assessment process for screeners, which includes both factual knowledge and practical ability. After commencing screening, performance is measured regularly across a broad range of key performance indicators (KPIs) and individual performance is compared to others within the same unit and the region.

There has been long-standing interest in factors affecting performance in colonoscopy[122]. In particular, the correlation of higher adenoma detection rates (ADR) in medium risk patients with a reduction in the risk of interval cancer[36] has prompted special-interest in this performance metric. Numerous studies have been performed looking at methods of improving ADR, including increasing colonoscopic withdrawal times[37,38,117,123], position change during extubation[120], the use of hyoscine[111,112,124], chromendoscopy[125] and other novel endoscopic techniques.

Analysis of data from the Bowel Cancer Screening Programme (BCSP) has found that the vast majority of individuals perform above the prescribed minimum standards. However, as expected in any population, some individuals perform consistently higher than others, even within the already selected group of BCS colonoscopists. Amongst other metrics, there appears to be a twofold difference in the ADR and fourfold difference in the mean number of adenomas detected per patient. The reasons for these differences are unclear but are not accounted for by known factors affecting adenoma detection rate.

6.2 Aims

6.2.1 Primary outcomes

The primary outcome was to determine factors contributing towards expertise in screening colonoscopy

6.3 Hypothesis

It is hypothesised that previously undetermined human factors correlate with KPIs in screening colonoscopists.

6.4 Methods

Research into human factors lends itself to qualitative rather than quantitative analysis[126]. It is well recognised that in qualitative studies, using one single method of enquiry is likely to result in inadequate data collection, and using multiple methods is much more likely to produce an accurate representation of the important human factors in individual disciplines[127,128].

The study was therefore planned in several steps to ensure important human factors were captured.

Figure 6.1. Schematic of methodology to highlight factors responsible for high performance in colonoscopy

6.4.1 Literature search

A literature search on PUBMED was performed to establish the current state of theoretical and applied knowledge on performance in gastrointestinal endoscopy.

6.4.2 Attribute identification focus group

This semi-structured group discussion was designed to determine whether published themes were pertinent to expertise in endoscopy or whether other criteria should be included. This was based on a 'brainstorming' exercise aiming to identify what skills or behaviours experts in the field consider important for expert endoscopy. Participants were asked to identify skills that they considered to be relevant to high performance in endoscopy.

A group of endoscopists including BCS screeners from several centres and endoscopy staff (nurses, support staff) were asked to participate in an initial focus group to provide the transfer of the limited themes from those identified in the literature search to endoscopy.

For convenience, the focus group comprised staff from a single centre, including four BCS endoscopists, consultant gastroenterologists with a specialist interest in endoscopy, a nurse consultant, trainee gastroenterologists, endoscopy nurses, secretarial and administration staff. This sample was chosen as it encompassed a group of professionals with an interest in the subject topic.

The focus group lasted 45 minutes and key themes from the focus group were recorded, initially for discussion with the group, and used to create a list of factors thought to be important.

6.4.3 Rating task – initial iteration

This task was designed for two purposes. The first was to stratify the identified factors in terms of importance to determine those considered to be more integral to high-performance endoscopy than others. The second aim of this stage was to ascertain whether other factors not considered by the focus group had been omitted.

An independent group of 39 BCS endoscopists, none of whom were present during the initial focus group, were polled at a meeting for a training session. They were asked to rank the themes derived from the initial focus group, as well as providing suggestions as to any omitted themes.

This sample was chosen to try to validate the themes identified initially by overcoming institutional bias by the inclusion of different BCS endoscopists from across England from a variety of different units.

6.4.4 Semi-structured interviews

All BCS endoscopists currently practising colonoscopy in England were emailed directly by the national director of the Bowel Cancer Screening Programme, explaining the nature of the proposed research. They were asked to reply to the study lead if they did not wish to participate in the research. . An information sheet about the semi-structured interview process was provided. It was made clear that there was no obligation to take part in the research and that all information provided would be anonymised.

After an interval to permit any colonoscopists to withdraw, further email contact was made by inviting screeners to provide their contact details if they wished to contribute to the research. Those agreeing to participate responded by email and a subsequent telephone interview was scheduled, lasting up to 60 minutes. This was recorded, with consent (see Appendix 3).

The interviews were conducted in several parts. The first part was a case study based on the participant's experience of a time that required the skills of an expert endoscopist.

The interviewee was asked to recount a case in detail, describing their thoughts, decisions, actions and communications with colleagues. The case study was chosen as this methodology allows sometimes abstract concepts to be put into a real-life perspective[129], which facilitates exploration of the issues and it was envisaged that this would allow deeper understanding of participants' views of expertise[130].

The second part was the skill identification exercise. Endoscopists were asked directly about the skills they felt were important in making an expert endoscopist. The discussion was then expanded to ask about how skills are currently developed in training. Given feedback from the earlier stages in the research, questions were asked on the specific differences between the skills needed for diagnostic and therapeutic colonoscopy.

In previous studies of this type, sample sizes of 25 have been considered acceptable[131,132]. However, as gastrointestinal endoscopy is a more specialised subset of medical practice, it was anticipated that saturation point would occur at an earlier point. The project therefore

aimed to interview 20 consultants in total, but with the caveat that if a wide sample of responses were received, additional participants could be recruited.

6.4.5 Gap analysis

Gap analysis is the study of performance differences between individuals within a group to elicit the reasons behind these differences with the ultimate aim of improving performance[133]. Screeners were asked to identify factors that they believe would improve their own performance and the performance of others within the programme in order to identify a 'gap' in expertise. Specific questions were asked to facilitate the gap analysis, including questions pertaining to the individual's current performance level, anticipated future performance level, the means by which they aimed to bridge the gap between the two and any barriers to change.

6.4.6 Independent rating task – second iteration

The final part was the attribute identification and rating task, which were performed for a second time using the themes identified from the focus group and independent raters.

To avoid bias, only after the interview was complete were interviewees asked to perform an online rating task, by assigning an importance to each of the themes previously identified from the previous stages of the research. Each theme was rated from 1 (most important) to 5 (least important).

6.5 Ethical approval

All participants for the focus group and semi-structured interviews were volunteers, and all information gained during this process was anonymised. The study was evaluated by the local ethics department and deemed not to need formal ethical approval.

6.6 External review

The study proposal was peer reviewed by the Bowel Cancer Screening Programme Research Committee, who gave formal permission for the study to recruit colonoscopists from the Programme.

6.7 Data analysis

The semi-structured interviews were recorded and anonymised. The audio from the interviews was then edited, enhanced to improve the sound quality and then professionally transcribed.

The first two interviews were performed in conjunction with a psychologist trained in qualitative research to ensure the interviews were appropriately conducted. The psychologist gave feedback after each interview to improve the performance of the principal interviewer. Only after the psychologist was comfortable that the interviews were conducted to a high standard were the interviews conducted by a sole interviewer.

All interviews were coded using specialist software by the main investigators (QSR NVivo quantitative analysis software). To avoid investigator bias, the initial two transcripts were selected for double-coding by the independent psychologist.

Data extracted from transcripts was used to identify the thematic framework, index the key themes and then chart and map the results using the framework approach. A preliminary taxonomy of the themes related to expertise was developed by the lead investigator in collaboration with the psychologist experienced in qualitative data analysis. A sample of the transcribed interviews were re-coded according to the preliminary taxonomy. This method of cross-checking data has been successfully used in previous studies[134].

Questions specifically pertaining to the gap analysis were analysed separately.

6.8 Results

6.8.1 Literature search

There is very little literature defining characteristics of expert endoscopists. The vast majority of published material concerns training in endoscopy[135-138], with some limited literature on the assessment of technical endoscopic ability using simulators in experts[139], but often as a comparator to unskilled endoscopists. There are some data on the correlation between technical and non-technical skills in anaesthetists, but the relative importance of these attributes in endoscopy is unknown[140]. The studies are summarised in Table 6.1.

Table 6.1. Summary of literature search

Authors	Study type	Sample size	Results
Ferlitsch et al	Virtual reality endoscopic simulator assessment	13 beginners, 11 experts	Beginners improved with training to the same level as the experts
Haycock et al	Simulator versus patient-based training	36 novices	Simulator trained endoscopists demonstrated higher levels of skill
Grantcharov et	Simulator assessment to assess ability to differentiate expertise	8 experienced endoscopists, 10 intermediate endoscopists, 10 novices	The simulator was able to distinguish between more and less experienced endoscopists
Sedlack et al	Bovine simulator assessment	13 experienced endoscopists, 13 intermediate skill endoscopists, 13 novices	Each group performed better than the less experienced groups with respect to caecal intubation time
Koch et al	Simulator based assessment of expertise	26 novices, 23 experts	Experts were faster, used more force and caused more discomfort
Riem et al	Analysis of performance during simulated intra-operative cardiac arrest.	50 trainee anaesthetists	Technical and non-technical skills significantly correlated to each other

6.8.2 Attribute identification – focus group

A broad variety of factors were thought to be important by participants.

Technical ability ranked highly in participants' perception of experts. The ability to "do what other endoscopists couldn't" routinely do as well as the ability to "deal with the unexpected" were perceived as important characteristics. A focus on quality was deemed a defining characteristic by some, with several of the four BCS endoscopists commenting upon the importance of the adenoma detection rate in screening colonoscopy. Another theme

emerging from the group included how experts possessed greater experience than others in terms of numbers of cases completed.

One participant thought peer recognition was important, stating “I’d let them scope me.” How this recognition was achieved, whether self-declared or independently recognised by colleagues was discussed, with one endoscopist considering that true experts could be defined partly by their academic publication record.

Non-technical qualities of experts were also featured. Self-insight was also thought to be important with expert endoscopists’ knowledge of their own competence and awareness of their limits discussed. The relevance of judgement in difficult situations was another theme, especially in dealing with complications. The importance of good interactions with patients and staff were also considered by some to be essential characteristics of experts.

Some group members highlighted how different skills were relevant to diagnostic and therapeutic colonoscopy.

The themes were then summarised to encompass the comments that had been received by all participants (Table 6.2).

Table 6.2. Themes highlighted by focus group

Ability to deal with complications
Ability to tackle cases others won't
Academic publication record
Adenoma detection rate
Communication skills
Declaration of expertise by others
Inter-personal skills with staff
Lifetime experience
Low complication rates
Manner with patients
Self-declaration of expertise
Staying calm under pressure
Usage of novel endoscopic techniques

6.8.3 Rating task

In total, 36 responses were received from individual anonymous BCS endoscopists.

Each individual item was ranked 1 to 13 for both diagnostic and therapeutic colonoscopy, with a score of one relating to the item the endoscopists felt was most important and 13 the least. No additional themes were suggested not already included in the list derived from the focus group. The consensus views in order of importance are shown in Table 6.3.

Table 6.3. Ranked themes from Bowel Cancer Screeners

Rank	Diagnostic colonoscopy	Therapeutic colonoscopy
1	Low complication rates	Ability to deal with complications
2	Adenoma detection rate	Staying calm under pressure
3	Inter-personal skills with staff	Low complication rates
4	Communication skills	Communication skills
5	Manner with patients	Inter-personal skills with staff
6	Staying calm under pressure	Ability to tackle cases others won't
7	Lifetime experience	Manner with patients
8	Ability to deal with complications	Adenoma detection rate
9	Declaration of expertise by others	Lifetime experience
10	Ability to tackle cases others won't	Usage of novel endoscopic techniques
11	Usage of novel endoscopic techniques	Declaration of expertise by others
12	Self-declaration of expertise	Self-declaration of expertise
13	Academic publication record	Academic publication record

6.8.4 Semi-structured interviews

In total, 267 BCS endoscopists were emailed and invited to participate. There were 21 responses, a response rate of 7.9%. Of these respondents, 20 interviews were conducted during the study period. The participant demographics are listed in Table 6.4. The sample size was deemed adequate as after the first six interviews no new themes emerged. This methodology is recognised as consistent with previous work in this field[141].

Table 6.4. Participant demographics

Sex	Male	14
	Female	6
Background	Physician	14
	Surgeon	4
	Nurse	2

The principal emergent themes mentioned by participants relating to expertise are listed in Table 6.5.

Table 6.5. Overarching themes relating to expertise

Theme	Number of participants (total 20)
Technical skills	20
Previous experience	19
Judgement / decision-making	18
Communication	18
Teamwork	15
Resources	11
Leadership	8

6.8.4.1 Technical skills

Technical skills were mentioned by all interviewees in some guise. The majority of interviewees chose to discuss a case of difficult EMR as the scenario they felt required an expert. In response to questioning about the importance of technical ability, different facets were highlighted by endoscopists as important (Table 6.6).

Table 6.6. Technical skills considered important by experts

Skill	Number of participants
Endoscopic mucosal resection	20
Detection of polyps	17
Lifting polyps	14
High caecal intubation rate	11
Comfortable examination	7
Detection of cancer	3
Motor skills	3
Importance of hand-eye coordination	2
Speedy examination	1

The most common technical skill was endoscopic mucosal resection, often involving difficulty lifting a polyp. This was universally mentioned by all participants.

Whether diagnostic and therapeutic skills were different was contested. Some drew a distinction between diagnostic and therapeutic skills:

> *"I think you can distinguish ... there's the technical ability to get round the colon in an efficient, pain-free manner consistently, that's one set of skills, and then a second set of skills is the therapy, so the judgement of knowledge and then the endoscopic fine motor skills and so on to manipulate this and to remove the polyp safely."*
>
> *"I think there's clearly a bit of overlap between them but I think you can be a very competent diagnostic colonoscopist without being an expert therapeutic colonoscopist. So I think there are some attributes that make a therapeutic colonoscopist that aren't necessarily found in every diagnostic colonoscopist. I think attitude is important, attitude towards risk I think is hugely important, and being prepared to perhaps approach things with a more surgical mentality would be a feature of the most advanced expert therapeutic colonoscopists that wouldn't be seen in expert diagnostic colonoscopists."*

Others however felt therapeutic colonoscopy involved an evolution of the skills required for diagnostic procedures rather than being fundamentally different:

> *Different is the wrong word. You've got to be able to have all the diagnostic skills to do therapeutic skills because otherwise you can't get there. It's the foundations and the first step. You don't build the second floor without the first floor. You can build buildings without foundations, they'll fall down but you can do it, but you can't build a second floor without a first floor. And to do the therapeutic skills, which are more advanced, you've got to be able to do the therapeutic stuff first. You've got to walk before you can run.*

Interestingly**,** three participants reported during the interviews that they performed colonoscopy on their colleagues and rated this as a marker of their expertise and their technical proficiency.

6.8.4.2 Previous experience

The value of experience when attempting a case needing an expert was almost universally mentioned by interviewees (19/20). The number of cases interviewees had tackled during their lifetimes ranged from 2000 to "tens of thousands".

Experts said that they relied on their previous experience "completely" or "heavily". One endoscopist questioned the conscious value of the experience they had gained as they had been "having been doing this sort of thing for an awfully long period of time one probably takes it for granted".

The incremental value of training experience over the years was also felt to contribute positively towards performance and tackling more difficult lesions:

"The sheer number of polyps and sheer number of patients that one has scoped during the years puts you in a position to be able to take on the more difficult stuff that experts take on."

Becoming a bowel cancer screening endoscopist itself was also mentioned to impact positively upon individual performance:

"So I think when you start as a bowel cancer screening colonoscopist it's quite scary because polyps are much bigger than you're used to…but as you do more and more then your confidence grows and your skills improve. My skills, certainly in polypectomy, improved enormously when I started bowel cancer screening."

6.8.4.3 Judgement/ decision making

The role of good judgement in expert colonoscopy was mentioned as frequently as that of experience. The most common themes emerging are listed in Table 6.7.

Table 6.7. Decision making skills

Skills	Number of participants
Awareness of own limitations	11
Forward planning	6
Insight into own ability	6
Adequate knowledge	4
Awareness of alternative treatment options	4
Attention to detail	3
Following instinct / heuristics	2

Expressions such as "do I think I can do this?" were commonly encountered during the interviews. Expert colonoscopists often seemed to question themselves about whether the

current treatment was correct, including during procedures. A good example of this was described by a colonoscopist when performing a difficult EMR:

> *"And at each stage I was thinking, is it safe to proceed, is it safe to proceed? Would this man be better and safer if I stopped and put him through another pathway? Because this was a big polyp and frankly at every stage during this I was thinking, can I do this? Can anybody do it? Can somebody do it better than me? And that was my thought process throughout most of the management that I had to do with him."*

6.8.4.4 Communication/ teamwork

Teamwork and communication were rated highly by most interviewees. Factors such as non-verbal communication and the ability to predict instructions before being asked were considered strong features of a good team by six respondents.

One endoscopist said: "You almost catch them out of the corner of your eye going to get something and it's only when you ask and it's there waiting for you."

Another noted:

> *I did another EMR this morning as well and it was a case of – I'm putting the snare in, I'm saying, "Open" and even before I'm saying it, it's opening. I'm saying, "Close." Even before I'm saying it, they're saying, "It's closing." I'm saying – because I find that everybody says, "Oh, you should always close it by yourself."*

This aspect of staff working closely together and communicating efficiently yet often silently was a recurrent theme:

> *And the best you can say about a team is when the team works smoothly and nobody really notices the fact there's a team going on, because if you notice there's a team it's usually because somebody's done something you weren't*

expecting or hasn't done something you were expecting. If a team works smoothly nobody notices.

Clear communication with the nursing staff was highlighted by 8 interviewees. Half of the interviewees also emphasised communication with the patient, in terms of keeping comfortable (7/20), instilling confidence (3/20) and explaining the procedure (2/20).

6.8.4.5 Resources

Interviewees' view of the resources that were important to them fell into two broad categories.

The majority (13/20) mentioned staff as a key resource and "that the staff that are supporting you, your endoscopy assistant is someone who you're confident in."

The second category of important resources was additional equipment, such as snares, lifting solution and diathermy machines. A broad range of equipment was not deemed to be essential: "it doesn't need to be a very wide variety, it just needs to be the right things". Familiarity with the equipment was deemed crucial by 9/20 respondents, with availability of the correct equipment instilling confidence in colonoscopists. One endoscopist commented: "I would never attempt to perform this sort of procedure with the other bit of equipment that I'm less comfortable with."

6.8.5 Gap analysis

There were a wide variety of views about how individuals felt their expertise within the screening programme could be improved (Table 6.8). In addition, all interviewees were able to suggest perceived deficiencies in the Bowel Cancer Screening Programme overall.

The most common improvement suggested was the introduction of peer-learning or mentoring to help screeners tackle more challenging lesions. However, when asked about the importance of the gaps that were identified, the majority (14/20) felt that before committing to a program to address the perceived gaps, it was important to ascertain the clinical relevance. One endoscopist asked: "if it's just very subtle variations [in performance] then is it clinically significant?"

Table 6.8. Gap analysis

Gap	Number of participants
Mentoring programme / peer teaching	6
More rigorous assessment of polypectomy ability	4
Improve adenoma detection rates / differences in ADR between screeners	3
Assessment of polyps	2
Caecal intubation rate	2
Lack of systematic data collection	1
Relevance of withdrawal time	1
Undersedation causing discomfort	1
Variation in management of large polyps between different centres	1

6.8.6 Rating task – second iteration

19 of the 20 interviewees completed the online rating task, ranking each previously defined attribute from 1 (most important) to 5 (least important). The median scores given for each attribute for both diagnostic and therapeutic colonoscopy are shown in Table 6.9 below, in descending order of importance.

Table 6.9. Relative importance of predetermined themes by interviewees

	Diagnostic	Therapeutic
Low complication rates	1	1
Adenoma detection rate	1	2
Manner with patients	1	2
Ability to deal with complications	2	1
Communication skills	2	1
Inter-personal skills with staff	2	1
Staying calm under pressure	2	1
Lifetime experience	2	1.5
Ability to tackle cases others won't	2.5	2
Declaration of expertise by others	2.5	2
Self-declaration of expertise	3	3
Usage of novel endoscopic techniques	3	3
Academic publication record	4.5	4.5

This ranking was largely similar to the order identified in the first iteration with themes such as academic publication record, usage of novel endoscopic techniques, self-declaration of expertise and declaration of expertise by others appearing at the bottom of both lists.

6.9 Discussion

It is unsurprising that technical ability rates highly in each of the phases of this work. Colonoscopy is by its very nature a practical skill and without a certain degree of ability, safe, comfortable and effective colonoscopy is not possible. Experts in this study however rarely admitted to personally struggling with the technical aspects of colonoscopy, even when describing their challenging cases.

The differences in the perception of skills needed for diagnostic and therapeutic colonoscopy were interesting. Although some did view the procedures as entirely different, others took a more nuanced view and thought that proficiency in diagnostic procedures was the "foundation" for competent therapeutic colonoscopy. Interestingly however, when asked to rate the themes at the end of the interview, the scores given by colonoscopists in each of these two domains were largely similar (see Table 6.9).

Although technical ability was the most common theme identified, other non-technical skills appeared very frequently. Judgement, communication, teamwork and leadership were all integral parts of experts' views of qualities that they and other expert colonoscopists possessed. There was very much more questioning about whether the decisions being taken were correct at each stage of the procedure.

The relevance of non-technical skills was confirmed by the ratings given by interviewees in the second iteration of the scoring task. Out of the seven highest-ranked qualities scoring one or two, four were related to non-technical skills, including a good patient manner, communication skills, interpersonal skills with staff and staying calm under pressure.

These findings have not been shown to date in endoscopy with no published studies correlating endoscopic outcomes with non-technical skills. However, in other areas such as surgery, some studies have shown a correlation between non-technical performance and technical outcome[142]. When surgical teams were assessed for their non-technical abilities and number of mistakes made during laparoscopic cholecystectomy, it was found that there was a negative correlation between surgeons' situational awareness and their error rates[142].

Other studies have mirrored these findings with poorer non-technical skills associated with higher rates of technical errors in surgeons[143-147].

It is in some ways predictable that these findings could be translated through to endoscopy, especially as with more complex procedures performed by experts the endoscopy room increasingly takes on certain characteristics of the formal operating suite.

If this is the case, the challenge is to develop methods of training that can reliably imbue new endoscopists with these skills in a less haphazard way than the simple experiential learning of the past. All of the respondents in this study had performed thousands of procedures; indeed some stated that they had performed so many over decades that they had stopped counting altogether.

There is some evidence that non-technical skills training can improve surgical outcomes, although the effect size has been small[148]. In the aviation industry, crew resource management training has been embraced for several decades to improve the way in which rapidly changing teams work together. Even in this field, partly as a consequence of the low numbers of adverse events, the overall effect in improving safety is still controversial[149].

Team work was recognised as an important theme contributing towards expertise by most (15/20) respondents. It has been shown that effective teams have common characteristics including shared goals, behavioural norms, defined roles, flexible leadership, good communication, and common shared resources[150]. Although interviewees were heavily reliant on their individual teams for their own performance, no endoscopist mentioned how team performance could be improved as a whole. It has been shown that formal team training

can be more effective than the team-building that naturally occurs from individuals working collaboratively together[151,152].

In terms of improving the quality of the Bowel Cancer Screening Programme overall, there were myriad responses received yet interviewees were in definite agreement that interventions to facilitate change only occur if there was a clinical basis for doing so. Some endoscopists felt that the focus on key performance indicators was in and of itself not productive as the vast majority of screeners are acknowledged to perform exceptionally well.

The logical next step for this work is to formulate interventions that could improve technical and non-technical skills and then assess whether the desired effects are seen in clinical practice. This is likely to be difficult however as, as in other arenas both in and out of medicine, the influence of any intervention is likely to be small and the difference therefore difficult to measure and conclusively prove. One study has shown that a one day course training multidisciplinary endoscopy teams improved awareness of patient safety knowledge and attitudes[153], but whether this has an effect on real patient care remains to be seen.

Although this research was confined to the Bowel Cancer Screening Programme, it is likely that the findings can be translated into general clinical endoscopic practice. Studies have shown that regular feedback, particularly in regard to the adenoma detection rate, can in itself improve performance[154-156]. The importance of non-technical skills alongside those targeting technical performance metrics is being increasingly acknowledged[157]. It may be that a combination of training in technical and non-technical skills is the most effective way of improving expertise generally in endoscopists, although how this can be most efficiently achieved remains unclear, with several models hypothesized[157]. The introduction of a

national endoscopic database (NED) is likely to highlight local differences in performance with greater ease than in the past and may itself drive standards towards those achieved in the best-performing centres.

6.10 Strengths and limitations

This is a large study with several different methodologies used to ascertain features of expert endoscopy. Data was collected from several different sources independently. As the participants were all volunteers, and the response rate to the email invitation relatively low, there is a chance that the results are affected by selection bias. Of course, it would not have been practical to interview unwilling participants. However, the usage of several geographically different groups of endoscopists is likely to have counteracted this hypothetical issue. The study was designed to focus on screening colonoscopists and as such it may not be possible to extrapolate the results to wider endoscopic practice. It is conceivable that if the study had been broadened to include other endoscopists then other themes relevant to expert endoscopy may have been identified.

7 Conclusions and future directions

This thesis has evaluated colonoscopic polypectomy training both in the United Kingdom and internationally, as well as colonoscopic expert performance and sought to understand the nature of expertise, as it pertains to endoscopy.

The situation in the United Kingdom, with structured polypectomy assessment a mandatory part of all trainees portfolios, clearly leads the world. No other country mandates polypectomy training, with both numbers of lesions tackled and competency, in the same way. The widespread adoption of initially ancillary programmes such as Train the Trainers has become an integral part of the delivery of polypectomy and colonoscopy training in the UK. These programmes are being mirrored in many other countries keen to adopt the same standards as those found in the UK.

However it is also clear that there is still considerable progress to be made with respect to training both domestically and internationally. In the UK the infrastructure and guidelines for delivering training may exist, with JAG and the JETS e-portfolio playing a leading role, but the quality of training between sites and trainers is yet to be assessed. Future work could well identify the ideal modes and frequency of training in order to optimise training outcomes in an environment where resources to train are often constrained.

Internationally, most countries performing significant amounts of polypectomy rely on their trainees having performed a minimum number of colonoscopy procedures without any formal polypectomy assessment. Where guidelines do exist, such as in Australia where 30 snare polypectomies are recommended, there is no associated competency framework to assess skills. The international survey did reveal that the majority of trainers and trainees

were unaware of the DOPyS as a means to assess polypectomy: most respondents, when asked about guidelines for polypectomy training, referred to guidelines about the follow-up of patients who had undergone polypectomy.

The 2004 UK audit demonstrated clearly the impact of effects of a largely untrained cohort of endoscopists performing polypectomy, with significant morbidity from perforation and haemorrhage[15]. There is good reason to suspect that internationally in many places complications from polypectomy will mirror those found in this study. This is particularly likely to be the case given that the participants in the international polypectomy survey were enthusiasts, and hence more likely to be competent at colonoscopy as well as more interested in training. The survey was biased towards the developed world and although this is where the majority of polypectomy is performed, there is increasing interest in colorectal cancer screening in developing countries where the adoption of Western lifestyles has led to an increase in the rates of colorectal cancer. In resource poor settings, where the ability to manage complications may be limited by equipment and infrastructure, the safety of polypectomy becomes even more important.

In order to address these issues and effect meaningful change, clinical practice is likely to follow high-quality published data. The technically ideal study would be to compare polypectomy outcomes in terms of short, medium and long-term complications between those who have been trained to competence using the DOPyS as a framework and those who are untrained, thereby validating the DOPyS as a training tool. It is unlikely such a study would gain ethical approval.

A useful surrogate study would be to try to define the learning curve of polypectomy, if it is shown that training with the DOPyS improves outcomes. The learning curve has already been established for diagnostic colonoscopy insertion. The DOPyS has been validated in the assessment of polypectomy competence and so the logical next step will be to utilise this tool in those performing polypectomy at the beginning of their training and then for each subsequent case to establish how many cases the average trainee needs to perform before being deemed competent. As with other skills, highlighting a particular number of cases performed to assure competence is unlikely to be beneficial as the pace of skills acquisition varies between trainees. Publication of an evidence-based international consensus statement addressing with the number of procedures and the method of assessment would be the ideal way to address the inconsistency in training that exists worldwide.

It must not be forgotten that polypectomy cannot take place without the endoscopist detecting a polyp in the first instance. This thesis has shown significant variability in adenoma detection, even amongst experts, and this is likely to be exaggerated amongst those training in colonoscopy. The ideal number of cases before trainees are adept at detecting polyps has also yet to be determined and would prove clinically relevant work given the widespread introduction of screening programmes around the world. The introduction of the National Endoscopy Database may well assist in providing this information, in the same way that the JETS electronic database permitted accurate calculation of the average number of cases trainees need to perform to be able to insert the colonoscope to the caecum reliably[50].

A retrospective review of guidelines for colonoscopy competency reveals that before the advent of high quality data, international guidelines suggested fewer cases were necessary for trainees to be deemed competent than it is now apparent are needed. The same

guidelines however have embraced the latest data and adapted accordingly. It seems likely that publication of similar data for polypectomy would lead to subsequent change in this way.

The studies pertaining to experts practising screening colonoscopy were also insightful. The data clearly showed that passing factual and practical tests was associated with consistently high performance afterwards with only a tiny minority of endoscopists not achieving the prespecified minimum key performance indicators for participation in screening. Unfortunately no comparator data were available for the performance of those endoscopists who were not successful at passing accreditation but continued practising colonoscopy outside the screening programme. A study comparing these two groups would add further weight to the validity of the accreditation.

Most of the colonoscopists tested by the BCSP accreditation process would never have had a formal test of their colonoscopy abilities, as they were trained in an era before JAG, and many would have had minimal initial training before embarking on independent practice[15]. Given the changes in colonoscopy certification for trainees, in which a similar test is undertaken by all those now who wish to practice independently, future work may well address the question of the relevance of a further accreditation in those who have already been certified as competent by JAG to similar standards.

Analysis of the differences between colonoscopists within the BCSP did reveal significant differences in a wide variety of key performance indicators. The reasons for this are not clear but are likely to be multifactorial. With any technical task performed by human beings, performance differences are expected. The respective contributions of pure technical ability, non-technical skills, environment, workload and equipment are yet to be defined and this

would prove an interesting avenue to explore in the future. The ultimate goal would be to address whether those performing less well could undergo an intervention, whether that be mentoring in technical or non-technical skills, with a resultant improvement in performance.

The thoughts of screeners about the nature of expertise in colonoscopy would be useful in informing such work. The predominant themes highlighted were technical ability, the importance of experience and the value of teamwork. Non-technical skills were accorded a surprising amount of weight by interviewees, with judgment and communication also ranked highly both qualitatively during the interviews and quantitatively in the rating task. The synergy between technical and non-technical skills seems to be acknowledged as the cornerstone of true endoscopic expertise. These skills could be learned passively but in order to improve expertise in all domains – not just in cancer screening – current training may have to be modified to include targeted teaching on themes highlighted by this study.

This thesis has focused especially on the hard outcomes from colonoscopy: adenoma detection rates, caecal intubation rates and polypectomy proficiency. There are already published data focused on technical methods to improve each of these parameters. Exploring the effect of non-technical skills training on expert endoscopists has not been attempted and seems a logical next step to move this work forward to assess whether the abilities of those already performing at a high level can be further honed.

However, screening itself is a small proportion of the overall endoscopic workload but provides a convenient environment in which to test these hypotheses due to the homogeneity of the patient population and relatively narrow performance characteristics of the endoscopists. The ultimate goal would be to develop a strategy to efficiently inculcate

expertise into the wider population of endoscopists so all patients undergoing endoscopy experience the best possible outcomes.

Bibliography

1. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;362:1795-803.

2. Gavin DR, Valori RM, Anderson JT, Donnelly MT, Williams JG, Swarbrick ET. The national colonoscopy audit: a nationwide assessment of the quality and safety of colonoscopy in the UK. Gut 2013;62:242-9.

3. 2015. (Accessed 12 December, 2016, at http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer.)

4. Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. CA Cancer J Clin 2014;64:104-17.

5. Atkin W, Saunders B. Surveillance guidelines after removal of colorectal adenomatous polyps. Gut 2002;51:v6-v9.

6. Day DW, Morson BC. The adenoma-carcinoma sequence. Major Probl Pathol 1978;10:58-71.

7. Fearon ER. Molecular genetics of colorectal cancer. Annual Review of Pathology: Mechanisms of Disease 2011;6:479-507.

8. O'Brien MJ, Yang S, Mack C, et al. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. The American journal of surgical pathology 2006;30:1491-501.

9. Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol 2011;42:1-10.

10. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. Bol Oficina Sanit Panam 1968;65:281-393.

11. Eide TJ. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. Int J Cancer 1986;38:173-6.

12. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology 1990;98:371-9.

13. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;366:687-96.

14. Viiala C, Zimmerman M, Cullen D, Hoffman N. Complication rates of colonoscopy in an Australian teaching hospital environment. Intern Med J 2003;33:355-9.

15. Bowles CJ, Leicester R, Romaya C, Swarbrick E, Williams CB, Epstein O. A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow? Gut 2004;53:277-83.

16. Heldwein W, Dollhopf M, Rosch T, et al. The Munich Polypectomy Study (MUPS): prospective analysis of complications and risk factors in 4000 colonic snare polypectomies. Endoscopy 2005;37:1116-22.

17. Rosen L, Bub DS, Reed JF, 3rd, Nastasee SA. Hemorrhage following colonoscopic polypectomy. Dis Colon Rectum 1993;36:1126-31.

18. Sorbi D, Norton I, Conio M, Balm R, Zinsmeister A, Gostout CJ. Postpolypectomy lower GI bleeding: descriptive analysis. Gastrointest Endosc 2000;51:690-6.

19. Levin TR, Zhao W, Conell C, et al. Complications of colonoscopy in an integrated health care delivery system. Ann Intern Med 2006;145:880-6.

20. Rutter MD, Nickerson C, Rees CJ, Patnick J, Blanks RG. Risk factors for adverse events related to polypectomy in the English Bowel Cancer Screening Programme. Endoscopy 2014;46:90-7.

21. Warren JL, Klabunde CN, Mariotto AB, et al. Adverse events after outpatient colonoscopy in the Medicare population. Ann Intern Med 2009;150:849-57, W152.

22. Chukmaitov A, Bradley CJ, Dahman B, Siangphoe U, Bouhaidar D, Warren JL. Polypectomy techniques, endoscopist characteristics, and serious gastrointestinal adverse events. J Surg Oncol 2014.

23. Repici A, Hassan C, Vitetta E, et al. Safety of cold polypectomy for <10mm polyps at colonoscopy: a prospective multicenter study. Endoscopy 2012;44:27-31.

24. Metz AJ, Moss A, McLeod D, et al. A blinded comparison of the safety and efficacy of hot biopsy forceps electrocauterization and conventional snare polypectomy for diminutive colonic polypectomy in a porcine model. Gastrointest Endosc 2013;77:484-90.

25. Cotton PB, Connor P, McGee D, et al. Colonoscopy: practice variation among 69 hospital-based endoscopists. Gastrointest Endosc 2003;57:352-7.

26. Minoli G, Meucci G, Prada A, et al. Quality assurance and colonoscopy. Endoscopy 1999;31:522-7.

27. Leung K, Pinsky P, Laiyemo AO, Lanza E, Schatzkin A, Schoen RE. Ongoing colorectal cancer risk despite surveillance colonoscopy: the Polyp Prevention Trial Continued Follow-up Study. Gastrointest Endosc 2010;71:111-7.

28. Farrar WD, Sawhney MS, Nelson DB, Lederle FA, Bond JH. Colorectal cancers found after a complete colonoscopy. Clin Gastroenterol Hepatol 2006;4:1259-64.

29. Pabby A, Schoen RE, Weissfeld JL, et al. Analysis of colorectal cancer occurrence during surveillance colonoscopy in the dietary Polyp Prevention Trial. Gastrointest Endosc 2005;61:385-91.

30. Pohl H, Srivastava A, Bensen SP, et al. Incomplete polyp resection during colonoscopy-results of the complete adenoma resection (CARE) study. Gastroenterology 2013;144:74-80 e1.

31. Efthymiou M, Taylor AC, Desmond PV, Allen PB, Chen RY. Biopsy forceps is inadequate for the resection of diminutive polyps. Endoscopy 2011;43:312-6.

32. Rembacken B, Hassan C, Riemann JF, et al. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). Endoscopy 2012;44:957-68.

33. Rabeneck L, Paszat LF, Hilsden RJ, et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. Gastroenterology 2008;135:1899-906, 906 e1.

34. Holme O, Loberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. JAMA 2014;312:606-15.

35. Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. Association of colonoscopy and death from colorectal cancer. Ann Intern Med 2009;150:1-8.

36. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;370:1298-306.

37. Lee TJ, Blanks RG, Rees CJ, et al. Longer mean colonoscopy withdrawal time is associated with increased adenoma detection: evidence from the Bowel Cancer Screening Programme in England. Endoscopy 2013;45:20-6.

38. Moritz V, Bretthauer M, Ruud HK, et al. Withdrawal time as a quality indicator for colonoscopy - a nationwide analysis. Endoscopy 2012;44:476-81.

39. Singh H, Nugent Z, Demers AA, Bernstein CN. Rate and predictors of early/missed colorectal cancers after colonoscopy in Manitoba: a population-based study. Am J Gastroenterol 2010;105:2588-96.

40. Samadder NJ, Curtin K, Tuohy TM, et al. Characteristics of missed or interval colorectal cancer and patient survival: a population-based study. Gastroenterology 2014;146:950-60.

41. Lasisi F, Rex DK. Improving protection against proximal colon cancer by colonoscopy. Expert Rev Gastroenterol Hepatol 2011;5:745-54.

42. Haycock AV, Patel JH, Tekkis PP, Thomas-Gibson S. Evaluating changes in gastrointestinal endoscopy training over 5 years: closing the audit loop. Eur J Gastroenterol Hepatol 2010;22:368-73.

43. Spier BJ, Benson M, Pfau PR, Nelligan G, Lucey MR, Gaumnitz EA. Colonoscopy training in gastroenterology fellowships: determining competence. Gastrointest Endosc 2010;71:319-24.

44. Spier BJ, Durkin ET, Walker AJ, Foley E, Gaumnitz EA, Pfau PR. Surgical resident's training in colonoscopy: numbers, competency, and perceptions. Surg Endosc 2010;24:2556-61.

45. Church J, Oakley J, Milsom J, Strong S, Hull T. Colonoscopy training: the need for patience (patients). ANZ J Surg 2002;72:89-91.

46. Sedlack RE. Training to competency in colonoscopy: assessing and defining competency standards. Gastrointest Endosc 2011;74:355-66 e1-2.

47. Park HJ, Hong JH, Kim HS, et al. Predictive factors affecting cecal intubation failure in colonoscopy trainees. BMC Med Educ 2013;13:5.

48. Lee SH, Chung IK, Kim SJ, et al. An adequate level of training for technical competence in screening and diagnostic colonoscopy: a prospective multicenter evaluation of the learning curve. Gastrointestinal endoscopy 2008;67:683-9.

49. Wexner SD, Garbus JE, Singh JJ, Group SCSO. A prospective analysis of 13,580 colonoscopies. Reevaluation of credentialing guidelines. Surgical endoscopy 2001;15:251-61.

50. Ward ST, Mohammed MA, Walt R, Valori R, Ismail T, Dunckley P. An analysis of the learning curve to achieve competency at colonoscopy using the JETS database. Gut 2014.

51. Benson M, Lucey M, Pfau P. Training the competent colonoscopist. Gastroenterology 2014;147:708-9.

52. Gupta S, Anderson J, Bhandari P, et al. Development and validation of a novel method for assessing competency in polypectomy: direct observation of polypectomy skills. Gastrointest Endosc 2011;73:1232-9 e2.

53. Anderloni A, Jovani M, Hassan C, Repici A. Advances, problems, and complications of polypectomy. Clin Exp Gastroenterol 2014;7:285-96.

54. Peluso F, Goldner F. Follow-up of hot biopsy forceps treatment of diminutive colonic polyps. Gastrointest Endosc 1991;37:604-6.

55. Hewett DG. Colonoscopic polypectomy: current techniques and controversies. Gastroenterol Clin North Am 2013;42:443-58.

56. Fyock CJ, Draganov PV. Colonoscopic polypectomy and associated techniques. World J Gastroenterol 2010;16:3630-7.

57. Norton ID, Wang L, Levine SA, et al. Efficacy of colonic submucosal saline solution injection for the reduction of iatrogenic thermal injury. Gastrointest Endosc 2002;56:95-9.

58. Tanimoto MA, Guerrero ML, Morita Y, et al. Impact of formal training in endoscopic submucosal dissection for early gastrointestinal cancer: A systematic review and a meta-analysis. World J Gastrointest Endosc 2015;7:417-28.

59. Tsuji Y, Fujishiro M, Kodashima S, et al. Desirable training of endoscopic submucosal dissection: further spread worldwide. Ann Transl Med 2014;2:27.

60. Draganov PV, Coman RM, Gotoda T. Training for complex endoscopic procedures: how to incorporate endoscopic submucosal dissection skills in the West? Expert Rev Gastroenterol Hepatol 2014;8:119-21.

61. Coman RM, Gotoda T, Draganov PV. Training in endoscopic submucosal dissection. World J Gastrointest Endosc 2013;5:369-78.

62. Ansell J, Hurley JJ, Horwood J, et al. The Welsh Institute for Minimal Access Therapy colonoscopy suitcase has construct and concurrent validity for colonoscopic polypectomy skills training: a prospective, cross-sectional study. Gastrointest Endosc 2014;79:490-7.

63. Training C, Sedlack RE, Shami VM, et al. Colonoscopy core curriculum. Gastrointestinal endoscopy 2012;76:482-90.

64. D W. Colonoscopy: Procedural Skills (AAFP Position Paper). Am Fam Physician 2000;62:1177-82.

65. Principles of training in gastrointestinal endoscopy. From the ASGE. American Society for Gastrointestinal Endoscopy. Gastrointestinal endoscopy 1999;49:845-53.

66. Bittner JGt, Marks JM, Dunkin BJ, Richards WO, Onders RP, Mellinger JD. Resident training in flexible gastrointestinal endoscopy: a review of current issues and options. Journal of surgical education 2007;64:399-409.

67. Short MW, Kelly KM, Runser LA. Colonoscopy by a family physician: a case series demonstrating health care savings. Military medicine 2007;172:1089-92.

68. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. Med Educ 1983;17:165-71.

69. Rees CE, Cleland JA, Dennis A, Kelly N, Mattick K, Monrouxe LV. Supervised learning events in the foundation programme: a UK-wide narrative interview study. BMJ Open 2014;4:e005980.

70. Barton JR, Corbett S, van der Vleuten CP, English Bowel Cancer Screening P, Endoscopy UKJAGfG. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: assessing colonoscopic skills of senior endoscopists. Gastrointest Endosc 2012;75:591-7.

71. Delfino AE, Chandratilake M, Altermatt FR, Echevarria G. Validation and piloting of direct observation of practical skills tool to assess intubation in the Chilean context. Med Teach 2013;35:231-6.

72. Ansell J, Hurley JJ, Horwood J, et al. Can endoscopists accurately self-assess performance during simulated colonoscopic polypectomy? A prospective, cross-sectional study. Am J Surg 2014;207:32-8.

73. Shanteau J. Expertise and Decision Support. In: George Wright FB, ed. The Psychology of Experts An Alternative View1992:11-23.

74. Chi MTH, Glaser R, Farr MJ. The Nature of expertise. Hillsdale, N.J.: L. Erlbaum Associates; 1988.

75. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science 1974;185:1124-31.

76. Anderson JR. Cognitive psychology and its implications. 7th ed. New York: Worth Publishers; 2010.

77. Shanteau J. Competence in experts: The role of task characteristics. Organ Behav Hum Decis Process 1992;53:252-66.

78. Glaser R, Chi MT, Farr M. The nature of expertise: Lawrence Erlbaum Associates; 1988.

79. Zsambok CE, Klein GA. Naturalistic decision making. Mahwah, N.J.: L. Erlbaum Associates; 1997.

80. Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003;133:614-21.

81. Flin RH, O'Connor P, Crichton M. Safety at the sharp end : a guide to non-technical skills. Aldershot, England ; Burlington, VT: Ashgate; 2008.

82. Ekkelenkamp VE, Koch AD, Haringsma J, Kuipers EJ, De Man RA. Endoscopist-related factors contributing to high-quality colonoscopy: results of a Delphi survey. Perspect Med Educ 2014;3:31-40.

83. Haycock A, Woolf K, Thomas-Gibson S. PTU-008 Enhancing professional behaviour in gastrointestinal endoscopy: development of a behavioural marker tool for assessment of endoscopic non-technical skills. Gut 2010;59:A51.

84. Haycock AV, Patel JH, Tekkis PP, Thomas-Gibson S. Evaluating changes in gastrointestinal endoscopy training over 5 years: closing the audit loop. Eur J Gastroenterol Hepatol 2010;22:368-73.

85. Sroka G, Feldman LS, Vassiliou MC, Kaneva PA, Fayez R, Fried GM. Fundamentals of laparoscopic surgery simulator training to proficiency improves laparoscopic performance in the operating room-a randomized controlled trial. Am J Surg 2010;199:115-20.

86. Majumdar D, Patnick J, Nickerson C, Rutter MD. Analysis of Colorectal Polyps Detected in the English Nhs Bowel Cancer Screening Programme with Emphasis on Advanced Adenoma and Polyp Cancer Detected. Gut 2012;61:A67-A.

87. Gschwantler M, Kriwanek S, Langner E, et al. High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. Eur J Gastroenterol Hepatol 2002;14:183-8.

88. Morris A, Hewitt J, Roberts CM. Practical experience of using directly observed procedures, mini clinical evaluation examinations, and peer observation in pre-registration house officer (FY1) trainees. Postgrad Med J 2006;82:285-8.

89. Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. BMJ 2010;341:c5064.

90. Coleman M, Hanna G, Kennedy R. The National Training Programme for laparoscopic colorectal surgery in England: a new training paradigm. Colorectal Dis 2011;13:614-6.

91. Miskovic D, Wyles SM, Carter F, Coleman MG, Hanna GB. Development, validation and implementation of a monitoring tool for training in laparoscopic colorectal surgery in the English National Training Program. Surg Endosc 2011;25:1136-42.

92. Coleman MG, Hanna GB, Kennedy R, National Training Programme L. The National Training Programme for Laparoscopic Colorectal Surgery in England: a new training paradigm. Colorectal Dis 2011;13:614-6.

93. Burford B, Illing J, Kergon C, Morrow G, Livingston M. User perceptions of multi-source feedback tools for junior doctors. Med Educ 2010;44:165-76.

94. Lockyer J, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. Teach Learn Med 2003;15:168-74.

95. Sargeant JM, Mann KV, Ferrier SN, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. Acad Med 2003;78:S42-4.

96. Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. Med Educ 2012;46:1042-54.

97. Draganov PV, Chang M, Coman RM, Wagh MS, An Q, Gotoda T. Role of observation of live cases done by Japanese experts in the acquisition of ESD skills by a western endoscopist. World J Gastroenterol 2014;20:4675-80.

98. Herreros de Tejada A. ESD training: A challenging path to excellence. World J Gastrointest Endosc 2014;6:112-20.

99. Bialek A, Pertkiewicz J, Karpinska K, Marlicz W, Bielicki D, Starzynska T. Treatment of large colorectal neoplasms by endoscopic submucosal dissection: a European single-center study. Eur J Gastroenterol Hepatol 2014;26:607-15.

100. Berr F, Wagner A, Kiesslich T, Friesenbichler P, Neureiter D. Untutored learning curve to establish endoscopic submucosal dissection on competence level. Digestion 2014;89:184-93.

101. Shiga H, Endo K, Kuroha M, et al. Endoscopic submucosal dissection for colorectal neoplasia during the clinical learning curve. Surg Endosc 2014;28:2120-8.

102. Oda I, Suzuki H, Nonaka S, Yoshinaga S. Complications of gastric endoscopic submucosal dissection. Digestive Endoscopy 2013;25:71-8.

103. Toyonaga T, Man-i M, East JE, et al. 1,635 Endoscopic submucosal dissection cases in the esophagus, stomach, and colorectum: complication rates and long-term outcomes. Surg Endosc 2013;27:1000-8.

104. Lewicki P, Czyzewska M, Hoffman H. Unconscious acquisition of complex procedural knowledge. J Exp Psychol Learn Mem Cogn 1987;13:523.

105. Stadler MA. On learning complex procedural knowledge. J Exp Psychol Learn Mem Cogn 1989;15:1061.

106. Rittle-Johnson B, Alibali MW. Conceptual and procedural knowledge of mathematics: Does one lead to the other? J Educ Psychol 1999;91:175.

107. Lee TJ, Rutter MD, Blanks RG, et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. Gut 2012;61:1050-7.

108. Adler A, Wegscheider K, Lieberman D, et al. Factors determining the quality of screening colonoscopy: a prospective study on adenoma detection rates, from 12,134 examinations (Berlin colonoscopy project 3, BECOP-3). Gut 2013;62:236-41.

109. Balentine CJ, Naik AD, Robinson CN, et al. Association of High-Volume Hospitals With Greater Likelihood of Discharge to Home Following Colorectal Surgery. JAMA Surg 2014.

110. Burnett-Hartman AN, Passarelli MN, Adams SV, et al. Differences in epidemiologic risk factors for colorectal adenomas and serrated polyps by lesion severity and anatomical site. Am J Epidemiol 2013;177:625-37.

111. Corte C, Dahlenburg L, Selby W, et al. Hyoscine butylbromide administered at the cecum increases polyp detection: a randomized double-blind placebo-controlled trial. Endoscopy 2012;44:917-22.

112. de Brouwer EJ, Arbouw ME, van der Zwet WC, et al. Hyoscine N-butylbromide does not improve polyp detection during colonoscopy: a double-blind, randomized, placebo-controlled, clinical trial. Gastrointest Endosc 2012;75:835-40.

113. Munson GW, Van Norstrand MD, O'Donnell J J, Hammes NL, Francis DL. Intraprocedural evaluation of comfort for sedated outpatient upper endoscopy and colonoscopy: the La Crosse (WI) intra-endoscopy sedation comfort score. Gastroenterol Nurs 2011;34:296-301.

114. Ekkelenkamp VE, Dowler K, Valori RM, Dunckley P. Patient comfort and quality in colonoscopy. World J Gastroenterol 2013;19:2355-61.

115. Atkin W, Rogers P, Cardwell C, et al. Wide variation in adenoma detection rates at screening flexible sigmoidoscopy. Gastroenterology 2004;126:1247-56.

116. McClements PL, Madurasinghe V, Thomson CS, et al. Impact of the UK colorectal cancer screening pilot studies on incidence, stage distribution and mortality trends. Cancer Epidemiol 2012;36:e232-42.

117. Saritas U, Ustundag Y. A continuing debate on 6-minute withdrawal time as a quality indicator during colonoscopy. Endoscopy 2012;44:970; author reply

118. Sawhney MS, Cury MS, Neeman N, et al. Effect of institution-wide policy of colonoscopy withdrawal time > or = 7 minutes on polyp detection. Gastroenterology 2008;135:1892-8.

119. Lee RH, Tang RS, Muthusamy VR, et al. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). Gastrointest Endosc 2011;74:128-34.

120. Koksal AS, Kalkan IH, Torun S, et al. A simple method to improve adenoma detection rate during colonoscopy: altering patient position. Can J Gastroenterol 2013;27:509-12.

121. Corley DA, Jensen CD, Marks AR. Can we improve adenoma detection rates? A systematic review of intervention studies. Gastrointest Endosc 2011;74:656-65.

122. Rex DK, Bond JH, Winawer S, et al. Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the US Multi-Society Task Force on Colorectal Cancer. The American journal of gastroenterology 2002;97:1296-308.

123. Gellad ZF, Weiss DG, Ahnen DJ, Lieberman DA, Jackson GL, Provenzale D. Colonoscopy withdrawal time and risk of neoplasia at 5 years: results from VA Cooperative Studies Program 380. Am J Gastroenterol 2010;105:1746-52.

124. Rondonotti E, Radaelli F, Paggi S, et al. Hyoscine N-butylbromide for adenoma detection during colonoscopy: a randomized, double-blind, placebo-controlled study. Dig Liver Dis 2013;45:663-8.

125. Hlavaty T, Huorka M, Koller T, et al. Colorectal cancer screening in patients with ulcerative and Crohn's colitis with use of colonoscopy, chromoendoscopy and confocal endomicroscopy. Eur J Gastroenterol Hepatol 2011;23:680-9.

126. Arbous MS, Grobbee DE, van Kleef JW, Meursing AE. Dutch case-control study of anaesthesia-related morbidity and mortality. Rationale and methods. Anaesthesia 1998;53:162-8.

127. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999;34:1189.

128. Golafshani N. Understanding reliability and validity in qualitative research. The qualitative report 2003;8:597-606.

129. Kohlbacher F. The use of qualitative content analysis in case study research. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research; 2006.

130. Larkin M, Thompson A. Interpretative phenomenological analysis. Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners 2003:101-16.

131. Creswell JW. Qualitative enquiry and research design: Choosing among five approaches. 2007.

132. Mason M. Sample size and saturation in PhD studies using qualitative interviews. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research; 2010.

133. Norman GR, Shannon SI, Marrin ML. The need for needs assessment in continuing medical education. BMJ 2004;328:999-1001.

134. Pope C, Ziebland S, Mays N. Analysing qualitative data. Qualitative Research in Health Care, Third Edition 2007:63-81.

135. Ferlitsch A, Glauninger P, Gupper A, et al. Evaluation of a virtual endoscopy simulator for training in gastrointestinal endoscopy. Endoscopy 2002;34:698-702.

136. Haycock A, Koch AD, Familiari P, et al. Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. Gastrointest Endosc 2010;71:298-307.

137. Grantcharov TP, Carstensen L, Schulze S. Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. JSLS, Journal of the Society of Laparoendoscopic Surgeons 2005;9:130-3.

138. Sedlack RE, Baron TH, Downing SM, Schwartz AJ. Validation of a colonoscopy simulation model for skills assessment. Am J Gastroenterol 2007;102:64-74.

139. Koch A, Haringsma J, Schoon E, de Man R, Kuipers E. A second-generation virtual reality simulator for colonoscopy: validation and initial experience. Endoscopy 2008;40:735.

140. Riem N, Boet S, Bould M, Tavares W, Naik V. Do technical skills correlate with non-technical skills in crisis resource management: a simulation study. Br J Anaesth 2012:aes256.

141. Joffe H, Yardley L. 4 Content And Thematic Analysis. Research methods for clinical and health psychology 2004;56.

142. Mishra A, Catchpole K, Dale T, McCulloch P. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. Surg Endosc 2008;22:68-73.

143. Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room: analysis of skills and roles. Ann Surg 2008;247:699-706.

144. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. Qual Saf Health Care 2009;18:109-15.

145. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. Qual Saf Health Care 2009;18:104-8.

146. ElBardissi AW, Wiegmann DA, Henrickson S, Wadhera R, Sundt TM, 3rd. Identifying methods to improve heart surgery: an operative approach and strategy for implementation on an organizational level. Eur J Cardiothorac Surg 2008;34:1027-33.

147. Wiegmann DA, ElBardissi AW, Dearani JA, Daly RC, Sundt TM, 3rd. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. Surgery 2007;142:658-65.

148. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. Quality and Safety in Health Care 2009;18:109-15.

149. Salas E, Wilson KA, Burke CS, Wightman DC. Does crew resource management training work? An update, an extension, and some critical needs. Human Factors: The Journal of the Human Factors and Ergonomics Society 2006;48:392-412.

150. Clark PR. Teamwork: building healthier workplaces and providing safer patient care. Crit Care Nurs Q 2009;32:221-31.

151. Leedom DK, Simon R. Improving team coordination: a case for behavior-based training. Mil Psychol 1995;7:109.

152. Stout RJ, Salas E, Fowlkes JE. Enhancing teamwork in complex environments through team training. Group Dyn 1997;1:169.

153. Matharoo M, Haycock A, Sevdalis N, Thomas-Gibson S. Endoscopic non-technical skills team training: The next step in quality assurance of endoscopy training. World journal of gastroenterology: WJG 2014;20:17507.

154. Kahi CJ, Ballard D, Shah AS, Mears R, Johnson CS. Impact of a quarterly report card on colonoscopy quality measures. Gastrointest Endosc 2013;77:925-31.

155. Boroff E, Atia MA, Leighton JA, et al. Impact of Regular Feedback on Proximal Polyp-and Adenoma Detection Rates.

156. Chaptini L, Laine L. Can I improve my adenoma detection rate? J Clin Gastroenterol 2015;49:270-81.

157. Anderson J. The future of gastroenterology training: instruction in technical skills. Frontline Gastroenterology 2012;3:i13-i8.

Dissemination

Review articles

Bowel Cancer: Focus on Earlier Diagnosis
Kinesh Patel, Siwan Thomas-Gibson
Clinical Focus Primary Care 2012, 6 (3): 193–197

Colonic polyps and an update on the bowel cancer screening programme
Patel K, Marshall S, Thomas-Gibson S
Clinical Medicine 2012 Dec;12(6):580-2.

Is competency assessment at the specialist level achievable?
Kinesh Patel, Omar Faiz, Siwan Thomas-Gibson
Transl Gastrointest Cancer 2013;2(3):110-111

Refereed conference proceedings

The Impact of the Introduction of Formalized Polypectomy Assessment on Training in the United Kingdom
Digestive Disease Week, Orlando, USA *May 2013*

An Evaluation of Screening Colonoscopists' Performance after a Structured Accreditation Process
Digestive Disease Week, Orlando, USA *May 2013*

An Evaluation of Screening Colonoscopists' Performance after a Structured Accreditation Process
BSG Annual Meeting, Glasgow *June 2013*

Experience in Polypectomy Training and Assessment: an International Survey
ESGE, Prague *April 2015*

Oral presentations

The Impact of the Introduction of Formalised Polypectomy Assessment on Training in the United Kingdom
BSG Annual Meeting plenary session, Glasgow *June 2013*

Experience in polypectomy training and assessment: an international survey
Digestive Disorders Federation, London *June 2015*

Experience in polypectomy training and assessment: an international survey
United European Gastroenterology Week, Barcelona, Spain *October 2015*

Papers

The impact of the introduction of formalised polypectomy assessment in the United Kingdom *In press*

An international survey of polypectomy training and assessment *In press*

Factors defining expertise in screening colonoscopy *Submitted*

Prizes

BSG Plenary Session best abstract *June 2013*

BSG Trainees' section oral abstract prize *June 2015*

UEG Travel grant *October 2015*

Appendix 1 Colonoscopy DOPS assessment tool

Formative DOPS Assessment Form

Colonoscopy and Flexible Sigmoidoscopy

JAG
Joint Advisory Group on GI Endoscopy

Colonoscopist

Trainer / Peer

Date (DD/MM/YYYY)

Scale and Criteria Key

4	Highly skilled performance
3	Competent and safe throughout procedure, no uncorrected errors
2	Some standards not yet met, aspects to be improved, some errors uncorrected
1	Accepted standards not yet met, frequent errors uncorrected
n/a	Not applicable

■ Major Criteria ▫ Minor Criteria

Criteria	Score	Comments
Assessment, consent, communication		
■ Obtains informed consent using a structured approach ○ Satisfactory procedural information ○ Risk and complications explained ○ Co-morbidity ○ Sedation ○ Opportunity for questions		
■ Demonstrates respect for patient's views and dignity during the procedure		
■ Communicates clearly with patient, including outcome of procedure with appropriate management and follow up plan. Full endoscopy report.		
Safety and sedation		
■ Safe and secure IV access		
■ Gives appropriate dose of analgesia and sedation and ensures adequate oxygenation and monitoring of patient		
■ Demonstrates good communication with the nursing staff, including dosages and vital signs		
Endoscopic skills during insertion and procedure		
▫ Checks endoscope function before intubation		
▫ Performs PR		
■ Maintains luminal view / inserts in luminal direction		
■ Demonstrates awareness of patient's consciousness and pain during the procedure and takes appropriate action		
▫ Uses torque steering and control knobs appropriately		
▫ Uses distension, suction and lens washing appropriately		
■ Recognises and logically resolves loop formation		
▫ Uses position change and abdominal pressure to aid luminal views		
▫ Completes procedure in reasonable time		
Diagnostic and therapeutic ability		
■ Adequate mucosal visualisation		
■ Recognises caecal/desc. colon landmarks or incomplete examination		
■ Accurate identification and management of pathology		
■ Uses diathermy and therapeutic techniques appropriately and safely		
■ Recognises and manages complications appropriately		

Case Difficulty

Extremely easy	Fairly easy	Average	Fairly difficult	Very challenging
1	2	3	4	5

Learning objectives for next cases

Appendix 2 DOPyS assessment tool

Royal College of Physicians | JAG Joint Advisory Group on GI Endoscopy

DOPyS: Polypectomy Assessment Score Sheet

Date: _____

Assessor: _____

Endoscopist: _____

Case ID: _____

A separate sheet should be used for each case. Up to five polyps from one patient may be documents on the same DOPyS score sheet.

Scale and Criteria Key

4	Highly skilled performance
3	Competent and safe throughout procedure, no uncorrected errors
2	Some standards not yet met, aspects to be improved, some errors uncorrected
1	Accepted standards not yet met, frequent errors uncorrected
N/A	Not applicable/Not assessable

Generic	Polyp 1	Polyp 2	Polyp 3	Polyp 4	Polyp 5
Optimising view of / access to the polyp:					
1. Attempts to achieve optimal polyp position					
2. Optimises view by aspiration/insufflation/wash					
3. Determines full extent of lesion (+/- use of adjunctive techniques e.g. bubble breaker, NBI, dye spray etc) if appropriate					
4. Uses appropriate polypectomy technique (e.g. taking into account site in colon)					
5. Adjusts/stabilises scope position					
6. Checks all polypectomy equipment (forceps,snare,clips,loops) available					
7. Checks (or asks assistant to) snare closure prior to introduction into the scope					
8. Clear instructions to and utilisation of endoscopy staff					
9. Checks diathermy settings are appropriate					
10. Photo-documents pre and post polypectomy					
Stalked polyps: Generic, then					
11. Applies prophylactic haemostatic measures if deemed appropriate					
12. Selects appropriate snare size					
13. Directs snare accurately over polyp head					
14. Correctly selects en-bloc or piecemeal removal depending on size					
15. Advances snare sheath towards stalk as snare closed					
16. Places snare at appropriate position on the stalk					
17. Mobilises polyp to ensure appropriate amount of tissue is trapped within snare					
18. Applies appropriate degree of diathermy					
Sessile lesions / Endoscopic mucosal resection: Generic, then					
19. Adequate sub mucosal injection using appropriate injection technique, maintaining views					
20. Only proceeds if the lesion lifts adequately					
21. Selects appropriate snare size					
22. Directs snare accurately over the lesion					
23. Correctly selects en-bloc or piecemeal removal depending on size					
24. Appropriate positioning of snare over lesion as snare closed					
25. Ensures appropriate amount of tissue is trapped within snare					
26. Tents lesion gently away from the mucosa					
27. Uses cold snare technique or applies appropriate diathermy, as applicable					
28. Ensures adequate haemostasis prior to further resection					
Post polypectomy					
29. Examines remnant stalk/polyp base					
30. Identifies and appropriately treats residual polyp					
31. Identifies bleeding and performs adequate endoscopic haemostasis if appropriate					
32. Retrieves, or attempts retrieval of polyp					
33. Checks for retrieval of polyp					
34. Places tattoo competently, where appropriate					
Polyp size	mm	mm	mm	mm	mm
Polyp site: C/AC/TC/DC/SC/Rectum					
Overall Competency at Polypectomy: 4/3/2/1					
Comments:					

Appendix 3 Schedule for high performance semi-structured interview

OVERVIEW

Introduction

Description of project to interviewee

- Aim to identify factors affecting expertise in GI endoscopy
- Liaise anonymously with Bowel Cancer Screening to inform future training / intervention
- No funding declarations

Explanation of the use of digital recorder

- To avoid copious note-taking
- Will be transcribed, de-identified then deleted.
- Sound check

Format

Part 1: *Performance example. 20 mins*

Describe a case from your own endoscopic experience that you consider required an expert endoscopist.

Part 2: *Skill identification exercise. 15 mins*

Identification of important skills characteristic of an expert endoscopist.

Part 3: *Gap analysis. 15mins*

Description of current personal performance, future aims for performance and strategies to achieve future aims

Part 4: *Sorting task. 10 mins*

Rate a number of skills identified by others to deal with expertise in endoscopy

Focus of interview

- Interested in skills and behaviours
- No judgment about performance
- Explain that there are no right/wrong answers

Any questions?

DETAILED SCHEDULE

Personal information

- Endoscopic experience (approx)
- OGD
- Colon
- ERCP/other
- Years as a consultant
- Size of unit
- No of screeners within unit

Part 1: *Performance example*

You were asked to think of a case from your own experience in endoscopy that required an expert endoscopist.

I will ask you to walk through the case a number of times

1. A brief description of the case
2. I will repeat back the key aspects to check
3. Describe the case again in more detail to focus on the particular skills required

Please give as much information as possible, but not personal details about the patient or members of staff.

Description of case and development of timeline

I will now repeat the case back to you to check the details and identify the key management points

Repeat back case

I would now like you to go through the case and give a description of what you were thinking, decisions you had to make, communications with colleagues, planning the tasks etc. If I think anything is particularly important, I may ask you further questions to identify particular points.

Go through case

Specific points to identify:

- Do you consider yourself to be an expert?
- Do others consider you an expert?
- (Do you think an expert endoscopist would have handled it differently?)
- Would you approach the same patient in the same way?
- What teamwork issues arose during the case?
- How important was technical ability?
- Who was in the leadership role?
- What information were you using to make your decisions?
- Did you have previous experience for you to draw on?
- How did you decide what option to take? What factors affected your decision?
- What resources did you have to support you? How did you use them?
- Were there any communication issues specific to this case?

Additional questions:

1. Can you tell me why you picked this case? What particular features required an 'expert'?
2. How do you think someone with less experience i.e. a trainee, might have handled this situation?

Part 2: *Skill identification exercise*

What kinds of skills are important or make an expert endoscopist?

Think about the skills that make a truly expert screening endoscopist and what might distinguish an expert from a novice, with particular focus on screening colonoscopy.

Question: How do you think these skills are currently developed?

Question: How might a trainee gain these skills?

Question: Do you think there are any differences between the skills needed for diagnostic and therapeutic cases?

Part 3: *Gap analysis*

How do you think endoscopists perform in the BCSP overall?

How do you think you individually perform within the cohort of BCS screeners? (establish baseline performance awareness)

What do you think are the challenges for addressing existing performance gaps within the BCSP overall? (identification of known gaps across the BCSP)

Do you think these performance gaps are important?

What are your individual performance goals? (identification of target performance)

Obtain list of performance objectives

What resources would help you achieving those goals? (identification of action)

Identify actions related to each individual performance objective

What, if anything, hinders you attaining the goals? (identification of barriers)

Part 4: *Sorting task*

We have identified some factors thought to contribute to expertise in endoscopy.

There are no right or wrong answers. Please rank each of the skills in terms of importance to you when discussing endoscopic expertise in the sphere of bowel cancer screening.

Use a scale of 1-5

1= not important at all

2= slightly important

3= fairly important

4= quite important

5= very important