



OPEN ACCESS

ORIGINAL ARTICLE

An analysis of the learning curve to achieve competency at colonoscopy using the JETS database

Stephen Thomas Ward,¹ Mohammed A Mohammed,² Robert Walt,³ Roland Valori,⁴ Tariq Ismail,³ Paul Dunckley⁴

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2013-305973>).

For numbered affiliations see end of article

Correspondence to

S T Ward, Centre for Liver Research and NIHR Birmingham Biomedical Research Unit, Level 5 Institute for Biomedical Research, University of Birmingham, Vincent Drive, Birmingham B15 2TT, UK; drsteward@yahoo.com

Received 28 August 2013
Revised 4 December 2013
Accepted 8 January 2014
Published Online First
27 January 2014



Open Access
Scan to access more
free content



► <http://dx.doi.org/10.1136/gutjnl-2014-306765>
► <http://dx.doi.org/10.1136/gutjnl-2014-307242>



CrossMark

To cite: Ward ST, Mohammed MA, Walt R, et al. *Gut* 2014;**63**:1746–1754.

ABSTRACT

Objective The number of colonoscopies required to reach competency is not well established. The primary aim of this study was to determine the number of colonoscopies trainees need to perform to attain competency, defined by a caecal intubation rate (CIR) $\geq 90\%$. As competency depends on completion, we also investigated trainee factors that were associated with colonoscopy completion.

Design The Joint Advisory Group on GI Endoscopy in the UK has developed a trainee e-portfolio from which colonoscopy data were retrieved. Inclusion criteria were all trainees who had performed a total of ≥ 20 colonoscopies and had performed ≤ 50 colonoscopies prior to submission of data to the e-portfolio. The primary outcome measure was colonoscopy completion. The number of colonoscopies required to achieve CIR $\geq 90\%$ was calculated by the moving average method and learning curve cumulative summation (LC-Cusum) analysis. To determine factors which determine colonoscopy completion, a mixed effect logistic regression model was developed which allowed for nesting of patients within trainees and nesting of patients within hospitals, with various patient, trainee and training factors entered as fixed effects.

Results 297 trainees undertook 36 730 colonoscopies. By moving average analysis, the cohort of trainees reached a CIR of 90% at 233 procedures. By LC-Cusum analysis, 41% of trainees were competent after 200 procedures. Of the trainee factors, the number of colonoscopies, intensity of training and previous flexible sigmoidoscopy experience were significant factors associated with colonoscopy completion.

Conclusions This is the largest study to date investigating the number of procedures required to achieve competency in colonoscopy. The current training certification benchmark in the UK of 200 procedures does not appear to be an inappropriate minimum requirement. The LC-Cusum chart provides real time feedback on individual learning curves for trainees. The association of training intensity and flexible sigmoidoscopy experience with colonoscopy completion could be exploited in training programmes.

INTRODUCTION

Colonoscopy training programmes are required to assess the competency of trainees performing colonoscopies.^{1 2} Although competency can be assessed by a range of factors, including patient comfort, appropriate sedation, adenoma detection rate and polyp retrieval rate, the caecal intubation rate (CIR—the extent to which the colon is

Significance of this study

What is already known on this subject?

- The caecal intubation rate (CIR) of 90% is a standard in routine use in colonoscopy practice and is a key indicator of competency.
- The number of colonoscopies required to achieve a CIR of $\geq 90\%$ is not well established, yet a set number of procedures forms part of training certification criteria both in the UK and USA.
- Previous studies investigating the number of colonoscopies required to achieve competency have been limited to small groups of trainees, mostly from single centres.
- Cumulative summation (Cusum) charts are widely used in healthcare to monitor processes of care, but requirements of monitoring during training/learning periods are not adequately reflected in standard Cusum charts.

What are the new findings?

- Our study interrogated the Joint Advisory Group (JAG) e-portfolio which includes data from all training centres in the UK. This provided a unique opportunity to analyse colonoscopy data submitted by 297 trainees in the early stages of their training.
- Two statistical methods were employed to determine the number of colonoscopies required to achieve a CIR of $\geq 90\%$. By moving average analysis, the cohort of trainees reached a CIR of 90% at 233 procedures. By LC-Cusum analysis, 41% trainees were competent after 200 procedures.
- The association between training intensity and previous flexible sigmoidoscopy experience with colonoscopy completion has not been reported previously.

How might it impact on clinical practice in the foreseeable future?

- Our study lends support to the current colonoscopy certification criteria in the UK. The visual feedback provided by LC-Cusum charts could form part of a future composite measure designed to define competency.
- This large study of UK endoscopy trainees will allow future trainees to benchmark their competency progression against a UK national average.
- Trainees should aim to have periods of intense focused colonoscopy training to increase their rate of competency acquisition and not be concerned about gaps in their training.

examined) remains an important parameter³ and is the focus of this study. It is accepted that not all colonoscopies can be completed to the caecum for reasons including poor bowel preparation, diverticular disease, strictures and patient discomfort.⁴ Therefore, a target CIR of 90% has been set by the Joint Advisory Group (JAG) on Gastrointestinal Endoscopy in the UK.³ Stipulation of the target CIR is necessary but insufficient without specifying the minimum number of procedures over which the target CIR of $\geq 90\%$ is to be achieved—too few and we risk trainees being certified prematurely, too many and we risk delaying the certification of otherwise competent trainees. Nevertheless, the minimal number of procedures that need to be performed to reach this standard is not well established. The American Board of Surgery have recommended that a minimum number of 50 colonoscopies be performed during training⁵ whereas the UK JAG stipulates 200.⁶

Previous studies have tried to determine the number of colonoscopies required to achieve a CIR of $\geq 90\%$ by following groups of inexperienced trainees through their training at individual centres.^{7–15} These studies reported discordant results, with some suggesting that a CIR of $\geq 90\%$ can be achieved after 150 procedures^{7–9} while others stipulate a requirement for more than 275 procedures.^{14–15} These differences may be explained by the small cohorts of trainees assessed ($n=8–41$), variation in competency progression and quality of training. Our primary aim was to determine, more definitively, the minimum number of colonoscopies that should be performed during training to reach competency and thereby enable national training bodies to make more informed decisions on colonoscopy training.

JAG has developed an e-portfolio for trainees to record their endoscopic experience—the JAG endoscopy training system (JETS) e-portfolio¹⁶ which is used by all endoscopy training units in the UK. We made use of this large database to overcome the shortcomings associated with small sample sizes. The JETS e-portfolio allows all UK endoscopy trainees to demonstrate their endoscopic experience, and facilitates both formative and summative assessments. Use of the JETS e-portfolio is a prerequisite for JAG certification. The e-portfolio was rolled out in 2009 and contains procedure data for over 2000 trainees. The database also allows for the investigation of factors associated with colonoscopy completion—a secondary aim of this study.

Determination of the minimum number of colonoscopies can be undertaken using two statistical (graphical) approaches—the moving average method¹⁷ and the learning curve cumulative sum (LC-Cusum) method.¹⁸ We adopted both approaches. The moving average method produces an empirical learning curve for the cohort of trainees by taking data available from all trainees. The LC-Cusum analysis¹⁸ involves an adaptation of the cumulative sum method widely used in healthcare^{19–20} to the learning context. The LC-Cusum produces a learning curve for each trainee, and when this curve crosses a predefined threshold, the trainee is deemed to have reached an acceptable level of performance.

METHODS

Subjects

The JETS e-portfolio database was interrogated to retrieve all colonoscopy records from the date of the portfolio's release on 1 September 2009 to 5 December 2012. The following fields were retrieved: endoscopist identifier, date of procedure, patient age and gender, diagnosis at colonoscopy, extent of procedure and extent completed by trainee.

The database was used to return details on all its registered users who had entered colonoscopy data. The following fields

were retrieved: number of different trainers, completion of a JAG colonoscopy course, and flexible sigmoidoscopy and oesophago-gastric duodenoscopy experience. Some users had entered the number of baseline colonoscopies they had performed prior to submitting the data to the JAG e-portfolio, although this was not a mandatory requirement. This baseline number was categorised as 'none', '1–20', '21–50', '51–100' and 'more than 100'. All users were surveyed to determine their number of baseline colonoscopies, training posts and any breaks in their training. Two reminders were sent.

Trainees who had submitted fewer than 20 colonoscopies to the database were excluded from further analysis, as this was the minimum number required to calculate a moving average completion rate. To ensure that learning curves were assessed only for trainees in the early stage of their endoscopy training, all users with a baseline number >50 were excluded from further analyses.

Primary outcome measure

The primary outcome measure was colonoscopy completion. For a colonoscopy to be defined as complete, two criteria had to be met. Firstly, the colonoscopy had to be performed either with trainer observation but no physical assistance from the trainer or independently without trainer observation. Secondly, the extent of the colonoscopy performed by the trainee had to be classified as 'caecum', 'terminal ileum', 'ileo-colon anastomosis' or 'neoterminal ileum'.

Statistical analyses

Two approaches (moving average and LC-Cusum) are available to determine the minimum number of procedures, and we used both. We report the proportion of trainees deemed competent by both methods.

Moving average analysis

The mean CIR was calculated for each trainee over blocks of 20 procedures, mirroring the methods of a recent study.¹⁴ The mean CIR for all trainees was calculated and plotted against the number of procedures performed. A trainee was deemed competent by this method if the moving average CIR exceeded 90%.

LC-Cusum analysis

Calculation of LC-Cusum scores has been described in detail elsewhere.¹⁸ Unlike other Cusum techniques, the LC-Cusum is designed for the learning/training environment and so for its null hypothesis (H_0), the LC-Cusum assumes that the process under scrutiny is unacceptable (ie, learning phase) while the alternative hypothesis (H_1) is that the process is in control (ie, competent phase). The LC-Cusum considers each procedure in sequence one at a time and does not penalise the trainee for early unsuccessful procedures because it has a holding barrier at zero. When the LC-Cusum crosses a predefined threshold (h), we conclude that sufficient evidence has been accumulated to indicate that the procedure has been learned (ie, the trainee is now competent) and hence H_0 is rejected in favour of H_1 . When the LC-Cusum does not exceed the threshold, monitoring should continue because competency has not yet been demonstrated.

Application of the LC-Cusum technique (like any Cusum plot) requires setting of four parameters to determine threshold values: (1) p_0 the acceptable failure rate, (2) p_1 the unacceptable failure rate, (3) the probability of rejecting H_0 when it is true (type I error) and (4) the probability of rejecting H_1 when it is true (type II error). Choices of these four parameters vary with context and are often a matter of judgment and/or

consensus from experts.²¹ A p_0 value in our case is defined as 0.1, as this is the performance standard set by the JAG.⁶ Thus for a successfully completed colonoscopy, a trainee's LC-Cusum plot will rise by 0.1 units and drop by 0.9 units ($=1-0.1$) for an incomplete procedure. A p_1 value is usually set to twice p_0 ,^{22, 23} although we consider a range of values which reflect a spectrum of more or less plausible values (0.15, 0.2, 0.25, 0.3). Typical values for type I and type II errors are 10% and, unlike power calculations for clinical trials, in the context of performance monitoring they are typically set to be equal.²⁴ The LC-Cusum was designed in a similar manner to the two one-sided test procedure to test for equivalence in clinical trials.²⁵

Patient, trainee and training factors

A mixed effects logistic regression model was constructed using colonoscopy completion as the outcome variable to determine which factors were independently associated with completion. To accommodate the hierarchical nature of the data, trainees and training centres were entered into the model as random intercept terms to accommodate the clustering effects—procedures within trainees and patients within centres. The random effects were designed to ensure that trainees worked in more than one centre (ie, trainees were not nested in centres).

The following trainee and training factors were entered into the model as fixed effects: number of procedures, training intensity (number of procedures per month), trainee specialty (physician, surgeon, nurse endoscopist), breaks in training >6 months, baseline experience of ≥ 100 sigmoidoscopies, baseline experience of ≥ 100 oesophago-gastric duodenoscopies and attendance at a JAG approved colonoscopy course. The total number of colonoscopies was converted into an ordinal variable of blocks of 50 procedures in order to make the calculated ORs more clinically meaningful.

These factors were all available from the JETS e-portfolio database or survey, and were chosen as the number of procedures and baseline experience,⁷⁻¹⁵ training intensity²⁶ and breaks from training²⁷ have all previously been shown to have an effect on competency. Patient factors (age, gender and diagnosis on colonoscopy) were entered into the model as fixed effects. These factors have also been shown to influence colonoscopy completion rates outside of the training context.²⁸ Twenty-eight different diagnoses could be entered into the diagnosis field, and multiple diagnoses were permitted. We collapsed the diagnostic terms into five different diagnoses (colorectal cancer, polyps, inflammatory bowel disease, diverticular disease, stricture), which appear to be clinically meaningful when considering completion, as reported in previous studies.^{4, 28, 29} These five diagnoses were entered as binary (present/absent) covariates in the statistical model.

Training intensity is highly dependent on breaks from training, such as rotating into a training post where there is no opportunity to perform colonoscopy. Therefore, an adjusted training intensity was calculated by dividing the number of procedures performed by the time period of continuous training. Continuous training was defined as the time from start to end date of submitted data, minus any periods of 90 days or more during which no data were submitted. It was assumed that any breaks from training of <90 days were unlikely to be due to a change in training post.

A p value <0.05 was considered to be statistical significant. All p values and ORs are quoted from the mixed effects model.

All data computations, graphics and statistical analysis were performed using R (R foundation for Statistical Computing, Austria), and the mixed modelling used the lme4 package.³⁰

Exploring the p_1 parameter

Unlike the acceptable failure rate (p_0) of 0.10, there is no set standard for the p_1 parameter. LC-Cusum charts were derived using different values for the p_1 parameter in order to determine the effects of p_1 on defining competency.

It was thought that colonoscopists achieving an average CIR $\geq 90\%$ over a 12 month period ought to cross the upper control limit and therefore be deemed competent by LC-Cusum analysis. Therefore, to further aid assessment of the p_1 parameter, LC-Cusum analysis was performed using sequential colonoscopy data from 15 independent colonoscopists at the Queen Elizabeth Hospital, Birmingham (QEH) and the Gloucestershire Royal Hospital, Gloucester (GRH), by interrogating the endoscopy reporting software (Unisoft Medical Systems, Middlesex and SQLScope, Gloucestershire, UK). Prospectively collected data over a 1 year period from January 2012 to January 2013 were analysed.

Validation of data submitted to the JETS database

To assess, at least partially, whether the JETS database captured all colonoscopies performed by trainees, validation of the data submitted to JETS by a sample of trainees was performed. We were able to obtain colonoscopy data for all trainees from the local endoscopy reporting database at QEH over the most recent 2 year period. Of these trainees identified from the local database, some trainees also submitted data to JETS. The number of colonoscopies submitted to JETS was compared with the number of colonoscopies logged for the trainee in the local endoscopy reporting database over the same time period. We compared the number of procedures in the two data sources using a linear scatterplot and Pearson's correlation coefficient and a mean difference Bland-Altman plot.³¹

RESULTS

A total of 169 515 colonoscopy procedures entered by 1331 different users were retrieved from the JETS e-portfolio database (figure 1). These procedures were performed in 300 different endoscopy units throughout the UK. A total of 1165 users were registered as trainees. Baseline data had been recorded by 384 trainees in the JETS e-portfolio. All trainees were invited to complete the survey: 585 responded, giving a response rate of 50.2%, although only 460 responses returned colonoscopy baseline data. Baseline data were therefore available for 844 trainees.

Trainees who had submitted fewer than 20 colonoscopies to the database were excluded from further analysis, yielding 651 trainees. Trainees with a baseline number >50 procedures were then excluded, yielding the final total number of trainees for further analysis as 297. These trainees had performed a total of 36 730 colonoscopies from 255 different training units. The median number of training units per trainee was 3 (IQR 2-4, range 1-9).

Analysis of competence

The mean CIR by the moving average method for all trainees was plotted against colonoscopy number (figure 2). The mean CIR was 67.9%, 76.2% and 84.1% after 100, 150 and 200 procedures, respectively. The mean CIR reached the standard of 90% at 233 colonoscopies.

LC-Cusum analysis (with $p_0=0.10$, $p_1=0.20$) defined trainees as competent or yet to reach competence, according to whether the upper control limit was crossed or not. By this method, 36 trainees were competent with 261 trainees yet to

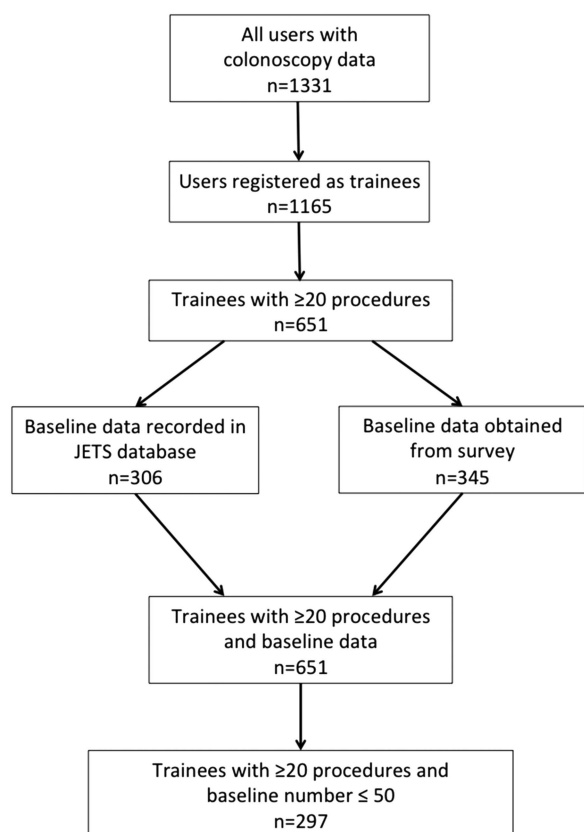


Figure 1 Study protocol. JETS, Joint Advisory Group endoscopy training system.

reach competency. The median number of procedures for the group of competent trainees to reach competency (ie, cross the upper control limit) was 171 (figure 3). It is important to emphasise that 171 procedures is not the median number of procedures required to achieve competency for all trainees. The

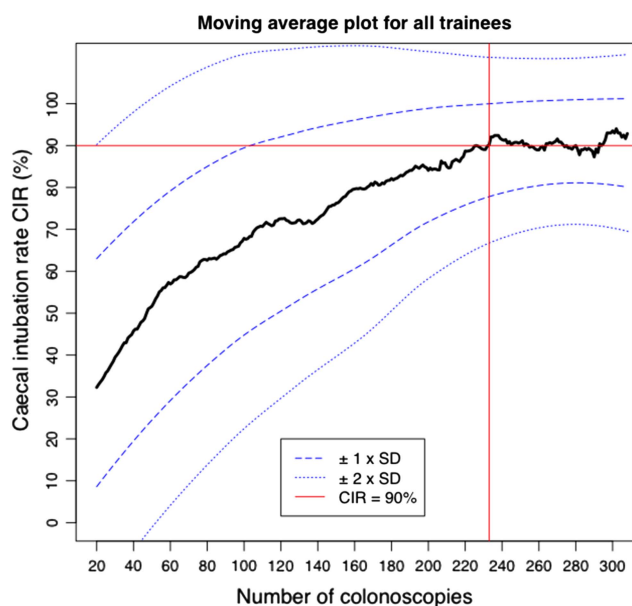


Figure 2 Mean moving average caecal intubation rate (CIR) against colonoscopy number (black line), plotted with 1 and 2 SDs of the mean. The mean CIR reached 90% at 233 colonoscopies.

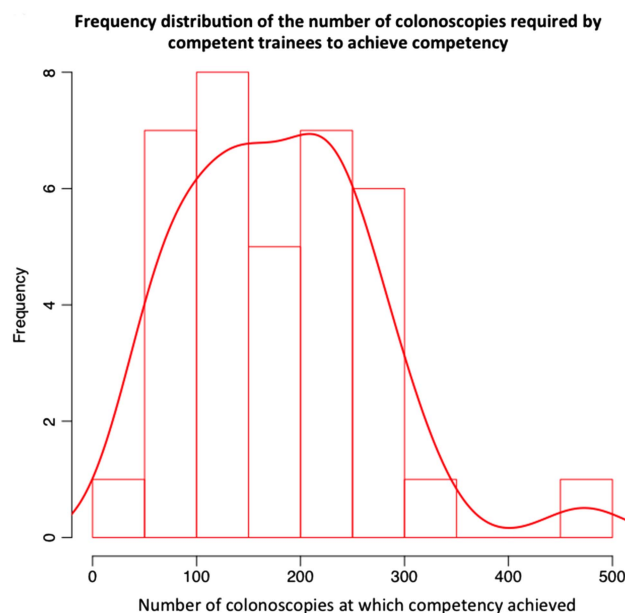


Figure 3 Median number of colonoscopies required to achieve competency for those trainees defined as competent by the learning curve cumulative summation (LC-Cusum) method.

majority of trainees (261) did not cross the upper control limit. Forty-nine trainees had performed 200 procedures or more: after 200 procedures, 21 trainees had attained competency and 30 trainees were yet to reach competency, giving a proportion of competent trainees of 41%. Thirty-three trainees had performed 250 or more procedures: after 250 procedures, 28 trainees had attained competency and nine trainees were yet to reach competency, a proportion of 76%.

Factors associated with colonoscopy completion

The results of the mixed effects logistic regression model is shown in table 1. In terms of trainee and training factors, an increasing number of procedures performed, a high training intensity, previous experience of more than 100 flexible sigmoidoscopies, a baseline level of colonoscopy experience of >20 procedures and being a surgical trainee were independently associated with colonoscopy completion. With regard to patient factors, female gender and age ≥ 65 years were significantly associated with colonoscopy incompleteness. The relationship between the proportion of colonoscopies completed against measured continuous variables is shown in figure 4. The proportion of completed colonoscopies increased with increasing number of procedures and training intensity. The proportion of completed colonoscopies decreased with increasing age above 60 years. A diagnosis was available for 83% colonoscopies, and submission of a diagnosis was independently associated with completion. The presence of colorectal cancer, diverticular disease and stricture were all significantly associated with incompleteness.

Adjustment of the p1 parameter

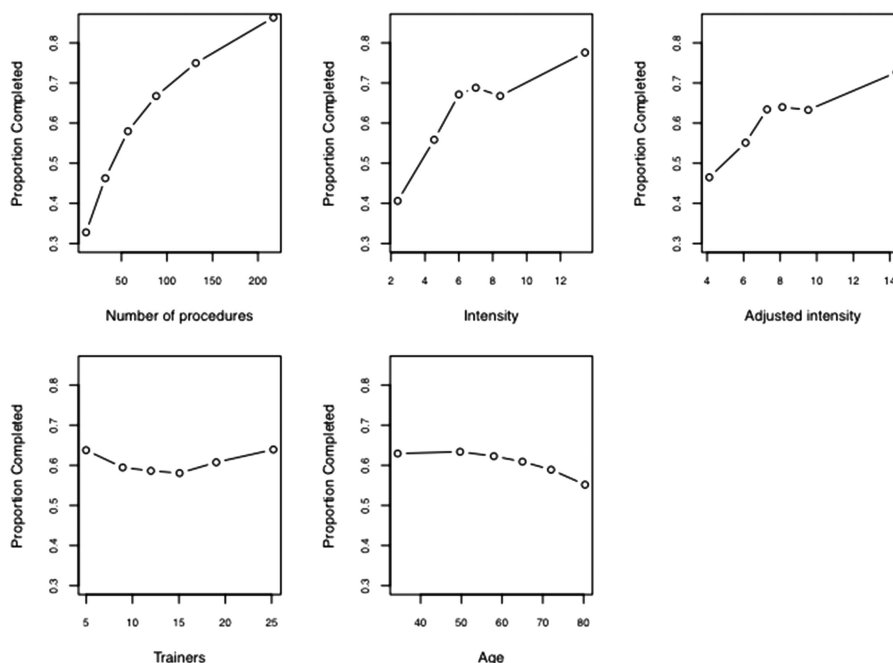
Different values of p1 yielded different results in terms of defining trainees as competent (see table 2). Adjustment of the p1 value from 0.20 to 0.25 increased the proportion of trainees that were deemed competent by the LC-Cusum method from 41% to 88%. An example of how different values of p1 affect the LC-Cusum chart for an individual trainee is shown in the online supplementary figure S1.

Table 1 Results of mixed effects model using colonoscopy completion as the outcome measure

Trainee and training factors	No of trainees (%)	OR (95% CIs)	p Value
No of procedures			
<50	65 (22)	Reference	
50–99	77 (26)	1.80 (1.50 to 2.20)	<0.01
100–149	55 (19)	4.20 (3.80 to 4.70)	<0.01
150–199	46 (15)	4.50 (4.0 to 5.00)	<0.01
200–249	28 (9.5)	4.60 (4.00 to 5.10)	<0.01
250–299	16 (5.5)	7.70 (6.90 to 8.40)	<0.01
300–349	9 (3)	6.40 (5.50 to 7.20)	<0.01
Adjusted intensity	297	1.10 (1.00 to 1.10)	0.034
No of trainers	297	0.99 (0.97 to 1.00)	0.68
Trainee specialty			
Gastroenterologist	167 (56)	Reference	
Gastrointestinal surgeon	97 (33)	1.60 (1.30 to 2.00)	0.01
Nurse endoscopist	32 (11)	1.30 (0.70 to 1.90)	0.40
General practitioner	1 (0.34)	1.20 (–1.10 to 3.40)	0.89
Breaks in training >6 months			
No	228 (77)	Reference	
Yes	69 (23)	1.10 (0.78 to 1.40)	0.56
Flexible sigmoidoscopy experience >100			
No	249 (84)	Reference	
Yes	24 (8.1)	2.40 (1.80 to 3.00)	<0.01
Unknown	24 (8.1)	1.00 (0.48 to 1.50)	0.98
OGD experience >100			
No	187 (63)	Reference	
Yes	69 (23)	1.20 (0.84 to 1.60)	0.30
Unknown	41 (14)	1.40 (0.94 to 1.90)	0.98
Completed colonoscopy course			
No	100 (34)	Reference	
Yes	197 (66)	0.88 (0.58 to 1.20)	0.38
Baseline level of experience			
None	166 (56)	Reference	
1–20	61 (21)	1.30 (0.85 to 1.70)	0.27
21–50	70 (24)	1.60 (1.20 to 2.10)	0.038
Source of baseline data			
JAG e-portfolio	92 (31)	Reference	
Survey	205 (69)	0.85 (0.42 to 1.30)	0.45
Patient factors	No of colonoscopies (%)	OR (95% CIs)	p Value
Age (years)			
<40	4 563 (12)	Reference	
40–64	17 934 (49)	1.10 (1.00 to 1.20)	0.063
>65	14 233 (39)	0.85 (0.77 to 0.94)	<0.01
Gender			
Male	19 450 (53)	Reference	
Female	17 280 (47)	0.64 (0.59 to 0.69)	<0.01
Diagnosis at colonoscopy (reference: not present)			
+Diagnosis available	30 458 (83)	1.80 (1.70 to 1.90)	<0.01
+Colorectal cancer	688 (2.3)	0.50 (0.33 to 0.68)	<0.01
+Polyps	8 126 (27)	1.10 (1.00 to 1.10)	0.072
+IBD	3 088 (10)	0.98 (0.89 to 1.10)	0.75
+Diverticular disease	5 455 (18)	0.86 (0.79 to 0.94)	<0.01
+Stricture	495 (1.6)	0.16 (–0.06 to 0.38)	<0.01
Random effects			
Trainee		1.00	
Hospital		0.88	

Percentages, ORs and p values given to 2 significant figures.
JAG, Joint Advisory Group; OGD, oesophagogastric duodenoscopy.

Figure 4 Scatterplot of the proportion of colonoscopies successfully completed against number of procedures performed, unadjusted intensity, adjusted intensity, number of different trainers and patient age.



It was thought that trainees achieving an average CIR of $\geq 90\%$ over a 12 month period ought to cross the upper control limit and therefore be deemed competent by LC-Cusum analysis. To test this, sequential colonoscopy data over 12 months from 15 independent practitioners (six from QEH, nine from GRH) were charted. All independent practitioners achieved a CIR of $\geq 90\%$ over the 12 month period and had performed in excess of 140 procedures (table 3). All practitioners crossed the upper control limit with p_1 set to values of ≥ 0.20 . When p_1 was < 0.20 , some independent practitioners failed to reach LC-Cusum defined competence. This suggests that the lower limit for p_1 used in LC-Cusum analysis of colonoscopy competency should be around 0.20.

Validation of data submitted to JETS

Twenty-eight non-independent colonoscopists performed colonoscopies over a 2 year time period (September 2010 to September 2012) at QEH. Ten of these 28 were registered and had submitted data to the JETS database. There was a high correlation between the number of procedures recorded in JETS and the number of procedures recorded in the local reporting software databases for these 10 trainees over this time period ($r^2=0.99$; $p<0.01$), but a Bland-Altman plot showed that the local number of procedures was on average three more than the JETS number (see online supplementary figure S2).

Table 2 Effect of different values of p_1 on the number of trainees reaching competency after 200 procedures

p_1 value	No of competent trainees	No of trainees yet to reach competency	Percentage of trainees who are competent (%)
0.15	0	54	0
0.20	21	30	41
0.25	70	10	88
0.30	116	5	96

DISCUSSION

This is the largest study investigating the colonoscopy learning curve to date and included all training centres in the UK. We have shown that a mean trainee CIR of 90%, calculated by moving average analysis, was reached at 233 procedures. By the LC-Cusum method, 41% of trainees were competent after 200 procedures and 76% were competent after 250 procedures. In order for a trainee to become provisionally certified as an independent colonoscopist in the UK, it is necessary for he/she to provide evidence that at least 200 procedures have been performed. For full certification (allowing fully independent practice), 300 procedures are required. Given our findings, these thresholds appear to be appropriate.

Advocates for lower numbers of procedures have argued that competency at performing a safe colonoscopy is more important than the absolute numbers performed.³² While this may be the case, the results from our study have shown that only a minority of trainees reach a CIR of 90% after 100 procedures. There is no apparent reason why trainees in the UK should develop colonoscopy skills at a slower rate than other countries, especially as there has been considerable investment and improvement in endoscopy training over the past decade.³³ This has resulted in a notable improvement in colonoscopy performance statistics in the UK.²⁸⁻³⁴ In comparison, European trainees performed 280 colonoscopies to reach a CIR of 90% by the moving average method in a Dutch study¹⁴ (as opposed to 233 in our study).

Other studies investigating the colonoscopy learning curve have been limited to small cohorts of between one and 24 trainees from individual centres, and have reported differing results.⁷⁻¹⁴ It has been shown that a CIR of $\geq 90\%$ can be achieved after only 140–150 procedures,⁷⁻⁹ although it is unclear what the absolute baseline level of colonoscopy experience was among these trainees. It is possible that these training centres have a particular interest in colonoscopy training, and it is noteworthy that the training intensity was reported as 23⁷ and 40⁹ colonoscopies per trainee per month—much higher than that found in our study. Our analysis has shown that intensity of training was a significant predictor of colonoscopy completion.

Table 3 Learning curve cumulative summation (LC-Cusum) analysis of competence of independent colonoscopists (numbered 1–15) using different values for p1

Colonoscopist	Hospital	N	Percentage complete	Practitioner crossed upper control limit (for different values of p1)							
				0.30	0.25	0.20	0.19	0.18	0.17	0.16	0.15
1	GRH	286	90	Yes	Yes	Yes	No	No	No	No	No
2	GRH	193	95	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
3	GRH	583	93	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4	GRH	165	90	Yes	Yes	Yes	Yes	No	No	No	No
5	GRH	301	91	Yes	Yes	Yes	Yes	Yes	Yes	No	No
6	GRH	413	93	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7	GRH	427	93	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8	GRH	217	92	Yes	Yes	Yes	Yes	Yes	No	No	No
9	GRH	521	91	Yes	Yes	Yes	Yes	Yes	Yes	No	No
10	QEH	413	92	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
11	QEH	287	94	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12	QEH	182	90	Yes	Yes	Yes	Yes	No	No	No	No
13	QEH	161	91	Yes	Yes	Yes	Yes	Yes	No	No	No
14	QEH	144	92	Yes	Yes	Yes	No	No	No	No	No
15	QEH	145	90	Yes	Yes	Yes	Yes	No	No	No	No
		Median n	Mean % complete	Percentage of practitioners who crossed upper control limit (for different values of P1)							
		286	92	100	100	100	87	67	53	40	27

GRH, Gloucestershire Royal Hospital; QEH, Queen Elizabeth Hospital.

While Cusum analysis of the colonoscopy learning curve was first reported in 1991,¹⁰ few studies have made use of this methodology in the context of assessing colonoscopy competency.^{8 27 35} This is despite its widespread use in monitoring performance and outcomes throughout healthcare.^{24 36} We have demonstrated that trainees attain competency at different rates and therefore a method that provides near real time feedback on an individual's learning curve, namely the LC-Cusum chart, yields more information on progress than absolute numbers. The LC-Cusum approach can provide near real time visual assessment of the trainee's learning curve that can be fed back to the trainee, trainers and training bodies, and aid formative and summative assessments. Furthermore, the upper control limit provides an objective measure of when competency has been attained. The LC-Cusum method could also provide an objective assessment of the effectiveness of training interventions, such as the use of simulators, training the trainers and immersion training. However, the LC-Cusum is dependent on four parameters (p0, p1, type 1 and type 2 errors) and it falls to the training bodies to set these.²¹ There is no precedent for the value of p1 in the context of CIR monitoring. However, the value of p1 typically lies between two and five times the acceptable failure rate.^{22 23} Other studies making use of Cusum charts in the colonoscopy context have set the p1 value to 0.20,²³ and in our study, we have provided supporting evidence that 0.20 appears to be an appropriate minimum, although further work is required to determine the optimum value. Trainees who have been deemed competent by the LC-Cusum method will be monitored to determine the CIR over their subsequent 100 procedures. The p1 standard will be defined as the largest value of p1 that ensures a CIR of $\geq 90\%$ over the 100 procedures following LC-Cusum defined competence.

Moving average analysis avoids the problems of requiring parameters to be set, other than the CIR standard. However, a decision has to be made as to the number of procedures over which to average. In our study, we chose blocks of 20 so that

the findings of our study were directly comparable with those recently published.¹⁴ This chart does not provide any individual feedback, like the LC-Cusum chart. Even though a moving average chart can be plotted using individual data, it does not provide real time feedback in the same way as the LC-Cusum chart because each data point is based on the most recent 20 procedures. Indeed, a fortunate clustering of 18 out of 20 successful procedures could signal that competency has been reached even though longer term consistency may not yet have been demonstrated. Nonetheless, the moving average method was useful in showing the group performance for all trainees. The ± 2 SD boundaries show that typical individual moving average plots could fluctuate within quite wide boundaries by chance alone.

We have shown that colonoscopy completion was associated with the total number of procedures performed, the intensity of training and experience at flexible sigmoidoscopy. Intensive training in the form of a 1 week course has been shown to improve skills following the course, an effect that was sustained for 9 months.²⁶ Our study found that breaks of 6 months or more from colonoscopy training had no significant effect on completion. This is consistent with the findings from a study comparing Cusum chart slopes of trainees following breaks from performing colonoscopy, where only small, if any, changes to the slope were noted.²⁷ These results suggest that trainees should focus on maximising their exposure to colonoscopy during their endoscopy training while being less concerned about prolonged gaps from their training.

Our analysis showed that being a surgical trainee was associated with successful colonoscopy completion. All trainees are by definition in the early stages of their endoscopy training, and it was therefore surprising that a trainee's chosen speciality should affect their ability to complete a colonoscopy. Colonic resection is reportedly associated with increased completion rates,^{29 37} and it is possible that higher completion rates in surgical trainees may relate to an increased proportion of patients

having had a prior bowel resection on surgical endoscopy lists compared with gastroenterology lists. We were unable to address this possibility as a history of colonic resection was not recorded in the JETS database.

We have demonstrated that patient age, gender and diagnosis were also associated with successful colonoscopy completion. This is consistent with other studies that have investigated factors affecting colonoscopy completion rates, albeit outside of the training context. Advancing age and female gender have been shown to be associated with incomplete colonoscopy by independent practitioners,^{37–39} as well as the presence of diverticular disease.²⁹

An important training factor is the use of a magnetic imager, as this has shown to result in improved performance by trainees⁴⁰ in terms of higher completion rates in a shorter time period. Unfortunately, the use of a magnetic imager was not routinely recorded in the JETS database, and we were therefore unable to assess the effects of this factor on completion.

One of the major limitations of this study is that colonoscopy completion was dependent on self-reported data. Trainees are required to accept an invitation from a hospital to be able to submit procedure data. The hospital endoscopy training lead and JETS administrator then have access to a trainee's data. When a trainee submits an e-portfolio for JAG certification, the training lead is required to validate the content of the e-portfolio. This allows for some validation of the data submitted. Furthermore, there was a strong correlation between trainee submitted data to the JETS e-portfolio and the local endoscopic reporting system, with small under-reporting, although this analysis was only possible using data from a small number of trainees (n=10).

Our study focused on CIR as one indicator of competence in performing colonoscopy. Our view is that it is the most important parameter in the context of supervised training. There are factors, in addition to CIR, that are critical to high quality colonoscopy: adenoma detection rate >10%; polyp recovery >90%; diagnostic colorectal biopsies for persistent diarrhoea in 100% cases; and appropriate tattooing. All of these measures assess the 'performance' of a colonoscopist. However, these factors are strongly influenced by the trainer who is present with the trainee at the time of the procedure. In the future, it is likely that a composite measure (including more than one measure of competency) will provide a more complete picture of colonoscopy performance. An alternative approach is to assess 'competencies' using competency frameworks such as the Mayo Colonoscopy Skills Assessment Tool (MSCAT)¹⁵ and the Direct Observation of Procedural Skills (DOPS).^{26–41} Trainers and assessors use these instruments to assess trainee competencies both to support training (formatively) and to determine overall competence (summatively). Although useful, the validity and reliability of these instruments to determine when a trainee is competent to practice independently are yet to be determined.^{42–43} Ultimately, it seems prudent to use a combination of assessment of performance and competency to assess fitness to practice independently.

In conclusion, this is the largest study to date investigating the number of procedures required to achieve competency in colonoscopy. Two statistical methods have been used, demonstrating that only 41% of trainees were competent after 200 procedures and that a mean CIR of 90% was reached after 233 procedures. Trainees accomplish competency at different rates, and the use of visual real time feedback tools, such as the LC-Cusum chart, gives more information on a trainee's progress than absolute numbers. Intensity of colonoscopy training and previous flexible

sigmoidoscopy experience were independently associated with attaining competency, and it is possible that these could be exploited by training programmes. Trainees may wish to start learning flexible sigmoidoscopy before colonoscopy training and should aim to have periods of intense focused colonoscopy training, not being overly concerned about gaps in training.

Author affiliations

- ¹Centre for Liver Research and NIHR Birmingham Biomedical Research Unit, Level 5 Institute for Biomedical Research, University of Birmingham, Birmingham, UK
²School of Health Studies, University of Bradford, Bradford, UK
³Department of Gastroenterology and GI Surgery, Queen Elizabeth Hospital, Birmingham, UK
⁴Department of Gastroenterology, Gloucestershire Royal Hospital, Gloucester, UK

Acknowledgements The authors would like to acknowledge the help of Kevin Dowler who retrieved data from the JETS e-portfolio for use in this study, and Dr D J Biau for advising us on the implementation of LC-Cusum using R. We would also like to acknowledge all of the trainees who responded to the survey.

Contributors Study concept: STW, MAM and PD. Study design: STW. Data acquisition: PD, RW and TI. Data/statistical analysis: STW and MAM. Manuscript preparation: STW. Manuscript review: PD, RW, TI and RV.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- NHS Cancer Screening Programmes, Great Britain, National Health Service. *Quality assurance guidelines for colonoscopy*. Sheffield: NHS Cancer Screening Programmes, 2011.
- 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 U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2002;97:1296–308.
- JAG Joint Advisory Group on GI Endoscopy. BSG Quality and Safety Indicators for Endoscopy. <http://www.thejag.org.uk/downloads%5CUnit%20Resources%5CBSG%20Quality%20and%20Safety%20Indicators.pdf> (accessed 7 Jan 2013).
- Brahmania M, Park J, Svarta S, et al. Incomplete colonoscopy: maximizing completion rates of gastroenterologists. *Can J Gastroenterol* 2012;26:589–92.
- Johna S, Klaristenfeld D. Surgery resident training in endoscopy: the saga continues. *Arch Surg* 2011;146:899–900.
- JAG—Certification Criteria. <http://www.thejag.org.uk/TrainingforEndoscopists/EndoscopyTrainees/TraineeCertification/CertificationCriteria.aspx> (accessed 16 Oct 2012).
- Lee S-H, Chung I-K, Kim S-J, et al. An adequate level of training for technical competence in screening and diagnostic colonoscopy: a prospective multicenter evaluation of the learning curve. *Gastrointest Endosc* 2008;67:683–9.
- Selvasekar CR, Holubar SD, Pendlimari R, et al. Assessment of screening colonoscopy competency in colon and rectal surgery fellows: a single institution experience. *J Surg Res* 2012;174:e17–23.
- Spier BJ, Benson M, Pfau PR, et al. Colonoscopy training in gastroenterology fellowships: determining competence. *Gastrointest Endosc* 2010;71:319–24.
- Parry BR, Williams SM. Competency and the colonoscopist: a learning curve. *Aust NZ J Surg* 1991;61:419–22.
- Tassios PS, Ladas SD, Grammenos I, et al. Acquisition of competence in colonoscopy: the learning curve of trainees. *Endoscopy* 1999;31:702–6.
- Chung JI, Kim N, Um MS, et al. Learning curves for colonoscopy: a prospective evaluation of gastroenterology fellows at a single center. *Gut Liver* 2010;4:31–5.
- Marshall JB. Technical proficiency of trainees performing colonoscopy: a learning curve. *Gastrointest Endosc* 1995;42:287–91.
- Koch AD, Haringsma J, Schoon EJ, et al. Competence measurement during colonoscopy training: the use of self-assessment of performance measures. *Am J Gastroenterol* 2012;107:971–5.
- Sedlack RE. Training to competency in colonoscopy: assessing and defining competency standards. *Gastrointest Endosc* 2011;74:355–366.e2.
- Mehta T, Dowler K, McKaig BC, et al. Development and roll out of the JETS e-portfolio: a web based electronic portfolio for endoscopists. *Frontline Gastroenterol* 2011;2:35–42.

- 17 Diggle P. *Time series: a biostatistical introduction*. Oxford University Press on Demand, 1990.
- 18 Biau DJ, Williams SM, Schlup MM, et al. Quantitative and individualized assessment of the learning curve using LC-CUSUM. *Br J Surg* 2008;95:925–9.
- 19 Sibanda T, Sibanda N. The CUSUM chart method as a tool for continuous monitoring of clinical outcomes using routinely collected data. *BMC Med Res Methodol* 2007;7:46.
- 20 Gombay E, Hussein AA, Steiner SH. Monitoring binary outcomes using risk-adjusted charts: a comparative study. *Stat Med* 2011;30:2815–26.
- 21 Yap C-H, Colson ME, Watters DA. Cumulative sum techniques for surgeons: a brief review. *ANZ J Surg* 2007;77:583–6.
- 22 Bolsin S, Colson M. The use of the Cusum technique in the assessment of trainee competence in new procedures. *Int J Qual Health Care* 2000;12:433–8.
- 23 Williams SM, Parry BR, Schlup MM. Quality control: an application of the cusum. *BMJ* 1992;304:1359–61.
- 24 Noye L. Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. *Interact CardioVasc Thorac Surg* 2009;9:494–9.
- 25 Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. *J Gen Intern Med* 2011;26:192–6.
- 26 Thomas-Gibson S, Bassett P, Suzuki N, et al. Intensive training over 5 days improves colonoscopy skills long-term. *Endoscopy* 2007;39:818–24.
- 27 Jorgensen JE, Elta GH, Stalburg CM, et al. Do breaks in gastroenterology fellow endoscopy training result in a decrement in competency in colonoscopy? *Gastrointest Endosc* 2013;78:503–9.
- 28 Bowles CJA, Leicester R, Romaya C, et al. 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.
- 29 Dafnis G, Granath F, Pahlman L, et al. Patient factors influencing the completion rate in colonoscopy. *Dig Liver Dis* 2005;37:113–8.
- 30 Bates D, Maechler M, Bolker B. *lme4: Linear mixed-effects models using Eigen and* 2012. <http://CRAN.R-project.org/package=lme4>
- 31 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- 32 ABS Statement on GI Endoscopy/American Board of Surgery. <http://www.absurgery.org/default.jsp?newsresponse> (accessed 7 Aug 2013).
- 33 Dunckley P, Elta G. Quality assurance of training. *Best Pract Res Clin Gastroenterol* 2011;25:397–407.
- 34 Gavin DR, Valori RM, Anderson JT, et al. The national colonoscopy audit: a nationwide assessment of the quality and safety of colonoscopy in the UK. *Gut* 2013;62:242–9.
- 35 Parry BR, Goh H-S. Quality control in colonoscopy: a Singapore perspective. *Int J Colorectal Dis* 1993;8:139–41.
- 36 Parry BR. Cusum: timely and trendy. *ANZ J Surg* 2007;77:523.
- 37 Tsai M-S, Su Y-H, Liang J-T, et al. Patient factors predicting the completion of sedation-free colonoscopy. *Hepatogastroenterology* 2008;55:1606–8.
- 38 Shah HA, Paszat LF, Saskin R, et al. Factors associated with incomplete colonoscopy: a population-based study. *Gastroenterology* 2007;132:2297–303.
- 39 Gupta M, Holub JL, Eisen G. Do indication and demographics for colonoscopy affect completion? A large national database evaluation. *Eur J Gastroenterol Hepatol* 2010;22:620–7.
- 40 Shah SG, Brooker JC, Williams CB, et al. Effect of magnetic endoscope imaging on colonoscopy performance: a randomised controlled trial. *Lancet* 2000;356:1718–22.
- 41 Barton JR, Corbett S, van der Vleuten CP, et al. 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.
- 42 Naeem N. Validity, reliability, feasibility, acceptability and educational impact of direct observation of procedural skills (DOPS). *J Coll Physicians Surg Pak* 2013;23:77–82.
- 43 Kogan JR, Holmboe ES, Hauer KE. Tools for direct observation and assessment of clinical skills of medical trainees: a systematic review. *JAMA* 2009;302:1316–26.