ORIGINAL PAPER

DOI: http://dx.doi.org/10.15403/jgld.2014.1121.281.nov

Training and Assessment in Flexible Sigmoidoscopy: using a Novel Direct Observation of Procedural Skills (DOPS) Assessment Tool

Keith Siau[1,2,3], James Crossley[4], Paul Dunckley[1,5], Gavin Johnson[1,6], Adam Haycock[7], John T. Anderson[1,5], Marietta Iacucci[3,8], Mark Feeney[1,9], on behalf of the Joint Advisory Group on Gastrointestinal Endoscopy (JAG)

1) Joint Advisory Group on Gastrointestinal Endoscopy, Royal College of Physicians, London; 2) Endoscopy Department, Dudley Group Hospitals NHS Foundation Trust, Dudley; 3) Institute of Immunology and Immunotherapy, NIHR Birmingham Biomedical Research Centre, University of Birmingham, Birmingham; 4) Academic Unit of Medical Education, University of Sheffield, Sheffield; 5) Department of Gastroenterology, Gloucestershire Hospitals NHSFT, Gloucestershire; 6) Department of Gastroenterology, University College London Hospitals NHSFT, London; 7) Wolfson Unit, St Mark's Hospital, London; 8) Institute of Translational Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham; 9) Department of Gastroenterology, Torbay and South Devon NHS Foundation Trust, Torquay, United Kingdom

Address for correspondence:
Keith Siau
Joint Advisory Group on Gastrointestinal Endoscopy, Royal College of Physicians, London, United Kingdom
keithsiau@nhs.net

Received: 21.12.2018
Accepted: 23.01.2019

ABSTRACT

Background & Aims: Data supporting milestone development during flexible sigmoidoscopy (FS) training are lacking. We aimed to present validity evidence for our formative direct observation of procedural skills (DOPS) assessment in FS, and use DOPS to establish competency benchmarks and define learning curves for a national training cohort.

Methods: This prospective UK-wide (211 centres) study included all FS formative DOPS assessments submitted to the national e-portfolio. Reliability was estimated from generalisability theory analysis. Item and global DOPS scores were correlated with lifetime procedure count to study learning curves, with competency benchmarks defined using contrasting groups analysis. Multivariable binary logistic regression was performed to identify independent predictors of DOPS competence.

Results: This analysis included 3,616 DOPS submitted for 468 trainees. From generalisability analysis, sources of overall competency score variance included: trainee ability (27%), assessor stringency (15%), assessor subjectivity attributable to the trainee (18%) and case-to-case variation (40%), which enabled the modelling of reliability estimates. The competency benchmark (mean DOPS score: 3.84) was achieved after 150-174 procedures. Across the cohort, competency development occurred in the order of: pre-procedural (50-74), non-technical (75-149), technical (125-174) and post-procedural (175-199) skills. Lifetime procedural count ($p<0.001$), case difficulty ($p<0.001$), and lifetime formative DOPS count ($p=0.001$) were independently associated with DOPS competence, but not trainee or assessor specialty.

Conclusion: Sigmoidoscopy DOPS can provide valid and reliable assessments of competency during training and can be used to chart competency development. Contrary to earlier studies, based on destination-orientated endpoints, overall competency in sigmoidoscopy was attained after 150 lifetime procedures.

Key words: competence – flexible sigmoidoscopy – formative assessment.

Abbreviations: DOPS: direct observation of procedural skills; DOPyS: the direct observation of polypectomy skills; EFA: exploratory factor analysis; ENTS: endoscopic non-technical skills; FS: flexible sigmoidoscopy; JAG: Joint Advisory Group on Gastrointestinal Endoscopy; JETS: JAG Endoscopy Training System; GP: general practitioner; NME: non-medical endoscopist.

INTRODUCTION

Flexible sigmoidoscopy (FS) is the third most commonly performed lower gastrointestinal examination in the UK, with over 500,000 procedures performed each year [1]. Flexible sigmoidoscopy is the first line investigation for anorectal symptoms and it is a primary colorectal cancer screening modality in England and Canada [2]. Over the last two decades, the focus on colonoscopy quality has catalysed reforms in colonoscopy training, instigating competency-based curricula, research into skills acquisition, and the development of a plethora of direct observational competency assessment tools to assure quality training [3]. In contrast, training in FS has received less prominence, as evidenced by the limited body of FS specific evidence. Given its service demands, there is an ongoing need for training in FS, and correspondingly, an onus on training programmes to ensure that training has been of sufficient quality to deliver a safe and effective FS workforce.

Competence during endoscopy training may be assessed using two main methods [3]: 1) indirectly, based on key

performance indicators [4, 5], or 2) directly, with direct observational assessment tools [6]. The latter allows for a trainee's skills to be evaluated by an observing assessor. In this context, assessments are performed with the objective of complementing training by highlighting procedure-specific strengths and weaknesses. Serial formative assessments can provide an indication of a trainee's progress, direct performance enhancing feedback, and indicate readiness for summative assessment and unsupervised practice [7].

The formative direct observation of procedural skills (DOPS) in lower gastrointestinal endoscopy (Supplementary material) were developed by the Joint Advisory Group on Gastrointestinal Endoscopy (JAG) [8] for FS and colonoscopy [9], and may be used interchangeably for both procedures. Although validity evidence has previously been presented for colonoscopy [10], similar data are lacking in FS. DOPS assessments are completed by supervising trainers on the JAG Endoscopy Training System (JETS) e-portfolio, a web-based platform used by all UK endoscopy trainees to record training procedures [11]. DOPS were developed following multidisciplinary expert consultation, which deconstructs components in sequential order based on procedure and non-technical elements into items with accompanying descriptors, enveloped within item groupings (domains). All DOPS are rated on a supervision-based scale and measure generic endoscopic non-technical skills (ENTS) [7, 12]. The current FS DOPS comprises 24 items organised into five domains (7 pre-procedure; 8 procedure; 3 management of findings; 2 post-procedure; 4 ENTS) and is separately complemented by an overall DOPS rating.

Determining when and how specific endoscopic competencies are attained will be useful to trainees, trainers and training programmes. Competence-assessment tools such as DOPS are well-placed to evaluate this when studied across a large training cohort.

Despite the implementation of DOPS into the UK curriculum, validity evidence to support its use in FS training remain lacking. This study had the following aims and objectives: i) to assess validity and reliability of formative DOPS, ii) to use DOPS to study and benchmark competence development (learning curves), and iii) to identify factors associated with DOPS competence.

METHODS

Study Design

This was a UK-wide observational study of in-training formative DOPS assessments for all FS procedures submitted onto the JETS e-portfolio in the 18-month period between July 2016 and December 2017. Direct observation of procedural skills completed based on the outdated performance-based scoring scale were excluded, as these differ with regard to competence assessment [7]. Under JAG recommendations, the decision for DOPS is made prior to commencing a procedure in order to minimise case-selection bias.

Study Covariates

For each DOPS, systematic data extraction was performed on the following: individual item scores, case difficulty overall assessor rating, trainee and assessor identifiers, specialties and training seniority. The lifetime procedural count for sigmoidoscopy immediately preceding the DOPS assessment date were collected; these were electronically determined based on trainee-populated JETS e-portfolio procedures. As some trainees also trained in colonoscopy, the lifetime FS count also incorporated colonoscopy procedures.

Outcomes

The primary outcome studied was the overall DOPS rating, which was independently awarded to the DOPS items and scored on a 4-point ordinal categorical scale [9]. In order to facilitate analysis, each outcome was converted into numeric form, i.e. Score 1 (requiring maximal supervision), Score 2 (significant supervision), Score 3 (minimal supervision), Score 4 (competent without supervision). Secondary outcomes included the mean DOPS score, and scores for pre-procedural, technical, post-procedural and ENTS domains. These were scored on a 4-point scale, but items could be rated "not applicable" (N/A) if assessment was not possible. The N/A scores were excluded from item and domain-level analyses.

Statistical Analysis

Exploratory Factor Analysis

Exploratory factor analysis (EFA) determines whether the distribution of scores within DOPS suggests that underlying latent variables are being assessed. Exploratory factor analysis was performed using a principle axis factoring with a threshold of Eigenvalue=1 and Varimax rotation in order to extract positively correlated factors into main groupings [13].

Generalisability Theory

Reliability estimates were performed using generalisability theory [14, 15], a conceptual framework which applies variance component analysis to estimate the influence of key assessment variables on overall DOPS rating. In this instance, variance components included: trainee ability (across all assessors and cases); assessor stringency (across all trainees and cases); assessor subjectivity attributable to the trainee; and general error, most of which will be case-to-case variation. From this data, generalisability coefficients (G) can be calculated as a function of the number of cases and assessors. The generalisability coefficient is based on the same equation as a reliability coefficient (subject variance/[subject variance + error variance]) and ranges between 0 (no reliability) and 1 (total reliability). A coefficient of ≥0.70 is the generally accepted threshold for in-training assessments [14].

Contrasting Groups Method

Each DOPS was stratified according to a global competency rating, i.e. Scores 1-3 as non-competent and 4 as competent. Distributions for mean DOPS scores (domain and item-total) were calculated for competent and non-competent groups, and subjected to contrasting groups method analyses (Fig. 1) [16]. The intercept for the two distributions was used to generate competency thresholds, which enabled consequence analysis by estimating theoretical false positive (FP) and false negative (FN) rates of competence.

Competency Development

To estimate competency development (learning curves), mean scores were calculated for each item, domain, and global

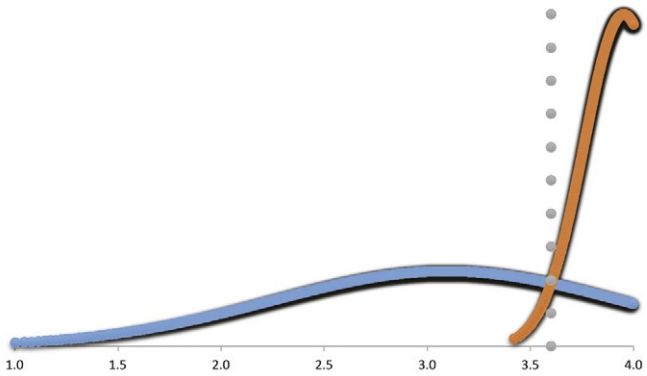

Fig. 1. Contrasting-groups method for ascertaining competency thresholds for a DOPS [16]. The x-axis refers to the mean DOPS score while the y-axis depicts the estimated proportion or frequency. In this example, distribution curves of mean DOPS scores were plotted according to non-competent (blue) and competent (orange) assessments based on the overall sigmoidoscopy DOPS score. The intersect of the two curves marks the ideal minimal threshold for the studied item or domain.

DOPS scores (mean item DOPS score and overall assessor rating) and analysed by strata of lifetime procedure count.

Multivariable Analyses (Generalised Estimating Equations)

Multivariable binary logistic regression analysis was performed to identify independent predictors of DOPS competence (overall DOPS score of 4). The generalised estimating equations (GEE) method using an autoregressive (AR1) structure was selected to account for the non-independence of procedures performed by the same trainee.

Statistical analyses were performed in SPSS (v24, Arkmont, NY: IBM Corp), with $p<0.05$ indicative of significance throughout.

RESULTS

Baseline Characteristics

In total, 3,616 FS DOPS submitted by 596 assessors (median DOPS per trainee: 2, IQR 1-6) from 211 UK training hospitals were included for analysis. These were completed for 468 trainees (median DOPS per trainee: 2, IQR 1-8) from gastroenterologist (N=189), gastrointestinal surgeon (N=113), non-medical endoscopist (NME; N=145) and general practitioner (GP; N=5) specialties. The median number of DOPS per trainee ranged across specialties ($p<0.001$) and was highest into the NME group (11, IQR 3-23), followed by GP (9, IQR 6-17), gastrointestinal surgeon (2, IQR 1-4) and gastroenterologist (2, IQR 1-3). The median lifetime procedure count also varied across specialty ($p<0.001$): gastroenterologist: 129 (IQR 44-245), gastrointestinal surgeon: 180 (IQR 89-303), NME: 172 (91-243) and GP: 64 (36-126).

The overall assessor DOPS ratings comprised: Score 1: 2.3%, Score 2: 9.2%, Score 3: 28.0% and Score 4: 60.5%. When analysed at item level, 91.6% of items were assessed, with 8.4% rated N/A. The 'N/A' ratings were most frequently applied to the items of: 'Complications': 63.3%, 'Sedation': 44.6% and 'Report Writing': 13.1%.

Factor Structure

Exploratory factor analysis (Table I) identified two positively correlated factors whose strongest loadings broadly correspond into 'pre-procedural' and 'procedural' groups. All factor loadings exceeded 0.4.

Table I. Exploratory factor analysis: rotated factor matrix revealing factor structure of DOPS.

DOPS Item	Factor	
	1	2
Indication		0.760
Risk		0.815
Confirms consent		0.882
Preparation		0.896
Equipment check		0.850
Monitoring		0.845
Sedation		0.711
Scope handling	0.801	
Tip control	0.804	
Air management	0.780	
Proactive problem solving	0.857	
Loop management	0.859	
Patient comfort	0.782	
Pace and progress	0.839	
Visualisation	0.845	
Recognition	0.799	
Management	0.827	
Complications	0.837	
Report writing	0.726	
Management plan	0.727	
Communication and teamwork	0.622	
Situation awareness	0.680	
Leadership	0.686	
Judgement and decision making	0.723	

Sources of Variance

Variance component analysis was performed to estimate the effect of key variables on the overall DOPS assessor rating. Sources of variance included: trainee ability (27%), assessor stringency (15%), assessor subjectivity attributable to the trainee (18%) and case-to-case variation (40%).

Reliability

Combining the variance estimates based on generalisability theory, the reliability of formative DOPS could be modelled on varying combinations of trainers and observations (Table II). Twelve observations (3 DOPS each from 4 different assessors) provided sufficient reliability to meet the reliability threshold of 0.70.

Establishing Competency Thresholds

Following contrasting groups analyses, competent and non-competent overall DOPS scores could be delineated based on

Table II. Reliability estimates (G-coefficients) for flexible sigmoidoscopy DOPS based on 1-8 trainers each observing 1-20 assessments. G-coefficients of 0.70+ based on assessor and assessment combinations (indicating sufficient reliability for in-training assessment) are shown in bold.

		Observations per Trainer							
		1	2	3	4	5	10	15	20
Trainers	1	0.27	0.34	0.37	0.39	0.40	0.43	0.44	0.44
	2	0.43	0.51	0.54	0.56	0.57	0.60	0.61	0.61
	3	0.53	0.61	0.64	0.66	0.67	0.69	**0.70**	**0.70**
	4	0.60	0.68	**0.70**	**0.72**	**0.73**	**0.75**	**0.76**	**0.76**
	5	0.65	**0.72**	**0.75**	**0.76**	**0.77**	**0.79**	**0.79**	**0.80**
	6	0.69	**0.76**	**0.78**	**0.79**	**0.80**	**0.82**	**0.82**	**0.83**
	7	**0.73**	**0.79**	**0.81**	**0.82**	**0.83**	**0.84**	**0.84**	**0.85**
	8	**0.75**	**0.81**	**0.83**	**0.84**	**0.84**	**0.86**	**0.86**	**0.86**

a mean DOPS score of 3.84, which led to FP and FN rates of 18.8% and 1.0%, respectively.

Sensitivity analysis was performed by excluding procedures performed by trainees with colonoscopy experience, i.e. lifetime colonoscopy count ≥20. This yielded similar results with a mean DOPS score threshold of 3.82 (FP: 16.9%, FN: 0.9%).

Competency Development in Trainees

Mean DOPS scores were presented by lifetime procedure count for each item (Table III) to chart competency development for specific skills across the cohort. At item-level, a mean score of 3.84 was set as a competency threshold. This showed that 50-74 procedures were required to attain pre-endoscopic competencies, 125-149 procedures for scope

Table III. Flexible sigmoidoscopy DOPS performance at item-level according to lifetime procedure count. The contrasting groups method was used to define the competency threshold for DOPS by comparing the mean DOPS score with the overall DOPS rating (see Figure 1). The scores marked in bold denote those which have passed the threshold (mean score of 3.84).

	Lifetime Procedure Count										
	≤ 24	25 - 49	50 - 74	75 - 99	100 - 124	125 - 149	150 - 174	175 - 199	200 - 224	225 - 249	250+
Indication	3.3	3.5	3.7	**3.8**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Risk	3.3	3.6	**3.8**	**3.8**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Confirms consent	3.5	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Preparation	3.4	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Equipment check	3.5	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Monitoring	3.5	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**	**4.0**	**4.0**
Sedation	3.3	3.7	3.7	**3.9**	**3.9**	**3.9**	**3.9**	**3.9**	**4.0**	**4.0**	**4.0**
Scope handling	2.7	3.2	3.5	3.5	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**3.9**	**4.0**
Tip control	2.6	3.1	3.4	3.4	3.6	3.7	**3.8**	**3.8**	**3.8**	**3.8**	**3.9**
Air management	2.7	3.2	3.5	3.6	3.7	**3.8**	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**
Proactive problem solving	2.4	3.0	3.2	3.3	3.5	3.6	3.7	**3.8**	**3.9**	**3.8**	**3.9**
Loop management	2.3	2.9	3.1	3.2	3.4	3.5	3.7	3.7	**3.8**	**3.8**	**3.9**
Patient comfort	3.0	3.4	3.5	3.6	3.8	3.7	**3.9**	**3.9**	**3.9**	**3.9**	**3.9**
Pace and progress	2.7	3.2	3.3	3.4	3.6	3.6	**3.8**	**3.8**	**3.8**	**3.8**	**3.9**
Visualisation	2.7	3.2	3.4	3.5	3.7	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**
Recognition	2.8	3.2	3.4	3.6	3.7	3.7	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**
Management	2.7	3.1	3.4	3.5	3.6	3.7	3.7	**3.8**	**3.8**	**3.9**	**3.9**
Complications	2.6	3.4	3.6	3.5	3.6	3.7	**3.9**	**3.9**	**3.9**	**3.9**	**4.0**
Report writing	2.8	3.2	3.4	3.3	3.6	3.7	3.7	**3.9**	**3.9**	**3.9**	**3.9**
Management plan	2.7	3.1	3.4	3.3	3.6	3.6	3.7	**3.8**	**3.8**	**3.9**	**3.9**
Communication and teamwork	3.3	3.5	3.7	**3.8**	**3.9**	**3.8**	**3.9**	**3.9**	**3.9**	**3.9**	**4.0**
Situation awareness	3.2	3.4	3.7	3.7	**3.8**	**3.8**	**3.9**	**3.9**	**4.0**	**3.9**	**4.0**
Leadership	3.1	3.4	3.6	3.6	**3.8**	**3.8**	**3.9**	**3.9**	**3.9**	**3.9**	**3.9**
Judgement and decision making	2.9	3.3	3.5	3.6	3.7	3.7	**3.8**	**3.8**	**3.9**	**3.9**	**3.9**
Mean Score	3.0	3.3	3.5	3.6	3.7	3.8	**3.8**	**3.9**	**3.9**	**3.9**	**4.0**

handling and air management, 150-174 for tip control, comfort, pace and progress, and visualisation. Loop management was the final skill to mature at 200-224 procedures. The post-procedural skills of management plan and report writing were acquired after 175-199 procedures. For generic endoscopic non-technical skills, competency development was observed in the order of communication and teamwork (75-99 procedures), situation awareness (100-124), leadership (125-149) followed by judgement and decision making (150-174). Using the mean score, 150-174 procedures were required to attain the DOPS competency benchmark.

Analyses were also performed at domain level (Fig. 2). This showed that competency development occurred in the order of: pre-procedure domain, ENTS domain, mean DOPS score, procedure domain, management domain, with post-procedure domain being the last to be acquired. These results followed those from item-level analyses.

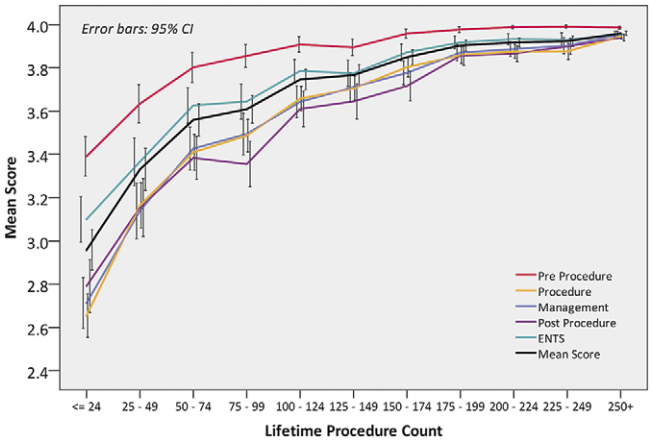

Fig. 2. Relationship between mean DOPS domain scores and lifetime procedural count.

Predictors of DOPS competency

On multivariable analysis (Table IV), lifetime procedural count (p<0.001), easier case difficulty (p<0.001), and lifetime flexible sigmoidoscopy DOPS count (p=0.001) were independently associated with overall DOPS competence, but not the trainee specialty (p=0.779), assessor specialty (p=0.332) and attendance of a basic skills course (p=0.652).

Table IV. Multivariable analysis of factors associated with competence (overall assessor score of 4) in formative FS DOPS. Generalised estimating equations were used to account for the non-independence of repeat DOPS by the same trainee, thereby providing analyses at trainee-level. *Excludes DOPS where data were unavailable **Includes lifetime colonoscopy counts. Bold p-values are significant at p<0.05.

Factor	N	Multivariable Odds Ratio	95% Confidence Interval	p-value
Specialty*				0.779
Gastroenterology	499 (14.3%)	REF		
GI Surgeon	440 (12.6%)	1.11	0.58 - 2.15	0.753
Non-medical endoscopist	2337 (67.0%)	0.80	0.46 - 1.40	0.804
General Practitioner	101 (2.9%)	1.07	0.36 - 3.22	0.905
Lifetime Procedural Count**				**<0.001**
<50	606 (17.4%)	REF		
50-99	496 (14.2%)	2.00	1.33 - 3.00	**0.001**
100-149	488 (14.0%)	2.82	1.76 - 4.56	**<0.001**
150-199	554 (15.9%)	9.29	5.55 - 15.56	**<0.001**
200-249	502 (14.4%)	12.57	7.15 - 22.10	**<0.001**
250+	842 (24.1%)	27.58	14.40 - 52.80	**<0.001**
Assessor Role				0.332
Gastroenterologist	1516 (43.5%)	REF		
Gastrointestinal Surgeon	721 (20.7%)	1.33	0.90 - 1.97	0.150
Non-medical endoscopist	1247 (35.8%)	1.26	0.90 - 1.74	0.177
Case Difficulty				**<0.001**
Easy	1231 (35.3%)	REF		
Moderate	2008 (57.6%)	0.51	0.41 - 0.64	**<0.001**
Complicated	249 (7.1%)	0.34	0.23 - 0.50	**<0.001**
Basic Skills Colonoscopy Course Attendance				
No	2762 (79.2%)	REF		
Yes	726 (20.8%)	1.11	0.71 - 1.73	0.652
Lifetime Flexible Sigmoidoscopy DOPS count				**0.001**
<10	1457 (41.8%)	REF		
10-19	778 (22.3%)	1.57	1.12 - 2.22	**0.009**
20+	1253 (35.9%)	2.67	1.61 - 4.44	**<0.001**

DISCUSSION

In the era of competency-based medical education, the role of continuous formative assessment during patient-based training is increasingly mandated by national accreditation bodies such as the National Accreditation System (NAS), American Society of Gastrointestinal Endoscopy (ASGE) and JAG [3]. In addition to the educational value provided to trainees, this approach enables trainers to monitor and verify progression of their trainee, and for training programmes to oversee competency development across the cohort of trainees, enable benchmarking, and to set milestones. Valid (well-grounded) and reliable (reproducible) formative assessment tools are therefore pivotal in facilitating competency-based endoscopy training.

This study, centred on FS DOPS for a pan-UK cohort of trainees, provides validity evidence for DOPS. The validity framework proposed by Messick, cites five sources of validity evidence in psychometric assessment: content, response, internal structure, relations to other variables and consequential validity [17]. The alignment between the EFA observed factors and the 'pre-procedural' and 'procedural' phases indicates expected internal structure (though this evidence alone is weak) [18]. Reliability is also intrinsic to psychometric validation; an assessment cannot be valid if measurements cannot be consistently reproduced. We have demonstrated that DOPS can be used to assess developing FS procedural skills with a level of reliability comparable to other workplace assessments. The learning curves analyses demonstrate the expected relationships with other variables, whilst the establishment of a pass-fail threshold confers consequential validity [18]. Content validity may be inferred from its expert multi-disciplinary implementation following an iterative DOPS process of task deconstruction, whereas response process validity may be surmised from survey data citing high trainee and assessor confidence in standards set by colonoscopy DOPS [10].

The literature on FS assessment is sparse. The landmark study from 1986 by Hawes et al. evaluated time to competency in 25 residents [19]. Trainees were graded on the percentage of mucosal visualisation and correct diagnoses, and on a six-point score for overall competence (1-3: non-competent, 4-6: competent), the authors concluded that 24-30 procedures were required to achieve competence in 85%+ of assessed procedures. However, in this small study, it is arguable that a single rating may have insufficient granularity to assess different competencies, and that ratings of competence may be affected by the nature of the scoring scale [7]. Sarker et al. [20] presented an assessment tool comprising 6 generic and 4 specific technical skills which demonstrated moderate scale reliability (Cronbach alpha: 0.79-0.81), with scores correlating with experience [20]. Thomas-Gibson et al. [21] evaluated the role of video assessment of FS extubations in FS screening practitioners, culminating in an assessment scoresheet which evaluated overall competence, and 5 specific performance areas: time spent viewing mucosa, re-examination of poorly viewed areas, suctioning of fluid pools, luminal distension, and lower rectal examination. This assessment tool showed good reliability (intraclass correlation coefficient 0.89) and was capable of discriminating between clusters of adenoma detection rates.

Contrary to earlier studies, which suggested that competency in FS might be achieved after 20-60 procedures, our data showed that overall competency was achieved in the cohort after 150 procedures. This discrepancy may be explained by several mechanisms, including: 1) the unsedated nature of FS performed in the UK, 2) the assessment of technical and non-technical skills, 3) the relative lack of access to pre-clinical simulation based teaching, 4) previous studies have based competency thresholds using sensitivities of 80%-90%, whereas our study, based on contrasting groups analysis, applied a mean DOPS score of 3.84 as a competency benchmark which provided a sensitivity (1-FN) of 99%. Reducing the competency benchmark to achieve a sensitivity of 80% would have led to a significantly higher FP rate, which risks rating non-competent trainees as being competent, at the potential disservice to patients.

The ASGE acknowledges that competence refers not only to technical achievement of performance standards of the examination (e.g. unaided in 80%-90% of examinations), but also to understanding the cognitive aspects of the procedure, including risks, benefits, and alternatives to the procedure [22]. Within the UK, training in endoscopy has evolved to ensure high quality patient care. This concept not only requires demonstration of effective and safe procedural skills, but also on the modern day attention to non-technical skills and post-procedural management. Trainees are increasingly expected to practice beyond the original endoscopy procedure and consider instigating the relevant necessary investigations to expedite the patient journey, which requires experience, medical knowledge and the understanding of a breadth of pathologies. Thus, the definition of competent within the UK may differ from other healthcare systems.

A limitation of real-world FS assessment is the lack of an evidence-based standard with regard to procedural extent. Within the literature, FS completion measures vary according to scope insertion depth and anatomical location, e.g. splenic flexure [20, 23, 24]. Within the UK, FS is performed without sedation and typically with enema preparation, which limits its extent [25]. Hence, trainees' skills may not be as robustly assessed compared to colonoscopy, which requires caecal intubation. Indeed, the majority of DOPS were rated as demonstrating overall competence. The relative lack of variance in the overall score attributable to the trainee (27%) was lower than the figure of 65% reported in senior endoscopists [26], which may explain the relatively high DOPS and assessor combinations required to achieve the in-training reliability threshold of 0.7. Data from multivariable analyses show that trainees were less likely to be rated competent in more complex cases. As competence in endoscopy requires independent and consistent performance across a range of case difficulties and contexts [3], the reliability models presented reflects this paradigm.

Several limitations specific to this study should be acknowledged. First, the observational study design enables inclusivity of analysis of formative assessments within a real-world training environment, but is inherently a source of bias. Trainees performed DOPS at variable intervals and frequencies,

the engagement of which differed by specialty. In response, JAG has attempted to improve standardisation by recommending a minimum of 1 DOPS per 10 training procedures. Second, lifetime procedure counts were calculated according to the number of JETS procedures logged by trainees. This has the potential for selection bias, which risks underestimating the procedural requirements in the learning curve analyses. Third, the majority of DOPS were performed for NME trainees, which may affect generalisability of data to other training cohorts comprising medical/surgical specialties. Non-medical endoscopists often do not have prior endoscopy experience (e.g. in pathology recognition and management), focus solely on FS training, and benefit from accelerated endoscopy training programmes, and hence, may require greater procedural numbers until competence is achieved. Despite this, NMEs appear to develop technical competencies on par with medical counterparts [27], which is supported by our multivariable analysis showing the lack of difference in DOPS competency outcome by trainee specialty. Fourth, we did not study specific polypectomy competencies, as these are measured on a separate instrument: the direct observation of polypectomy skills (DOPyS) [9]. Finally, as discussed above, variance analyses suggest a high error rate which impacts on the reliability estimates and the numbers of formative DOPS to accurately gauge competence. Future studies involving trainer assessment of videotaped procedures may be required to assess the inter-rater reliability of DOPS and to evaluate sources of DOPS variation.

CONCLUSION

This study provides evidence of validity and reliability in support of DOPS, and profiles competency development in a range of technical and non-technical competencies at item and domain level. Contrary to earlier studies based on destination-orientated endpoints, overall competency in FS was attained after 150 lifetime procedures. Our data may be of relevance to training programmes considering minimum procedural requirements and competency-based curricula, and for screening accreditation programmes centred on FS.

Conflicts of interests: K.S.: research fellow funded by JAG; J.C.: commissioned by JAG to produce generalisability theory analyses; P.D., G.J., A.H., J.T.A., M.F.: affiliated with JAG. No conflicts of interest for the other authors.

Authors' contributions: K.S. drafted the original manuscript and performed the literature review. K.S. and J.C. performed the statistical analyses. P.D., G.J., A.H., J.T.A. and M.F. aided in the development and implementation of the sigmoidoscopy DOPS and were collectively responsible for the study concept. J.C., P.D., G.J., A.H., J.T.A., M.I. and M.F. performed the critical review and enhanced the final version of the manuscript

Declarations: This paper presents independent research supported by the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Supplementary material: To access the supplementary material visit the online version of the *J Gastrointestin Liver Dis* at http://dx.doi.org/10.15403/jgld.2014.1121.281.nov

REFERENCES

1. Shenbagaraj L, Thomas-Gibson S, Stebbing J, et al. Endoscopy in 2017: a national survey of practice in the UK. Frontline Gastroenterol 2018;10:7-15. doi:10.1136/flgastro-2018-100970
2. Siau K, Yew AC, Ishaq S, et al. Colonoscopy conversion after flexible sigmoidoscopy screening: results from the UK Bowel Scope Screening Programme. Colorectal Dis 2018;20:502-508. doi: 10.1111/codi.13982
3. Siau K, Hawkes ND, Dunckley P. Training in Endoscopy. Curr Treat Options Gastroenterol 2018;16:345-361. doi:10.1007/s11938-018-0191-1
4. Siau K, Hodson J, Valori RM, Ward ST, Dunckley P. Performance indicators in colonoscopy after certification for independent practice: outcomes and predictors of competence. Gastrointest Endosc 2018 Aug 1. doi:10.1016/j.gie.2018.07.032
5. Ward ST, Hancox A, Mohammed MA, et al. The learning curve to achieve satisfactory completion rates in upper GI endoscopy: an analysis of a national training database. Gut 2017;66:1022-1033. doi:10.1136/gutjnl-2015-310443
6. Walsh CM. In-training gastrointestinal endoscopy competency assessment tools: Types of tools, validation and impact. Best Pract Res Clin Gastroenterol 2016;30:357-374. doi:10.1016/j.bpg.2016.04.001
7. Siau K, Dunckley P, Valori R, et al. Changes in scoring of Direct Observation of Procedural Skills (DOPS) forms and the impact on competence assessment. Endoscopy 2018;50:770-778. doi:10.1055/a-0576-6667
8. Siau K, Green JT, Hawkes ND, et al. Impact of the Joint Advisory Group on Gastrointestinal Endoscopy (JAG) on endoscopy services in the UK and beyond. Frontline Gastroenterol 2018 Nov 13. doi:10.1136/flgastro-2018-100969
9. Joint Advisory Group on Gastrointestinal Endoscopy. Formative DOPS: Colonoscopy and Flexible Sigmoidoscopy. 2016. Available from: https://www.thejag.org.uk/Downloads/DOPS%20forms%20(international%20and%20reference%20use%20only)/Formative%20DOPS_Colonoscopy%20and%20Flexible%20sigmoidoscopy.pdf
10. Barton JR, Corbett S, van der Vleuten CP; English Bowel Cancer Screening Programme; UK Joint Advisory Group for Gastrointestinal Endoscopy. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: assessing colonoscopic skills of senior endoscopists. Gastrointest Endosc 2012;75:591-597. doi:10.1016/j.gie.2011.09.053
11. Mehta T, Dowler K, McKaig BC, Valori RM, Dunckley P. Development and roll out of the JETS e-portfolio: a web based electronic portfolio for endoscopists. Frontline Gastroenterol 2011;2:35-42. doi:10.1136/fg.2010.003269
12. Hitchins CR, Metzner M, Edworthy J, Ward C. Non-technical skills and gastrointestinal endoscopy: a review of the literature. Frontline Gastroenterol 2018;9:129-134. doi:10.1136/flgastro-2016-100800
13. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014;51:511-521. doi:10.1016/j.ijnurstu.2013.10.005
14. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based

assessment scales. Med Educ 2011;45:560-569. doi:10.1111/j.1365-2923.2010.03913.x
15. Webb NM, Shavelson RJ, Haertel EH. 4 Reliability Coefficients and Generalizability Theory. In: Rao CR, Sinharay S, eds. Handbook of Statistics: Elsevier 2006:81-124.
16. Jørgensen M, Konge L, Subhi Y. Contrasting groups' standard setting for consequences analysis in validity studies: reporting considerations. Adv Simul (Lond) 2018;3:5. doi: 10.1186/s41077-018-0064-7
17. Messick S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. Am Psychol 1995;50:741-749. doi:10.1037/0003-066X.50.9.741
18. Downing SM. Reliability: on the reproducibility of assessment data. Med Educ 2004;38:1006-1012. doi:10.1111/j.1365-2929.2004.01932.x
19. Hawes R, Lehman GA, Hast J, et al. Training resident physicians in fiberoptic sigmoidoscopy. How many supervised examinations are required to achieve competence? Am J Med. 1986;80:465-70. doi:10.1016/0002-9343(86)90721-7
20. Sarker SK, Albrani T, Zaman A, Patel B. Procedural performance in gastrointestinal endoscopy: an assessment and self-appraisal tool. Am J Surg 2008;196:450-455. doi:10.1016/j.amjsurg.2007.10.024
21. Thomas-Gibson S, Rogers PA, Suzuki N, et al. Development of a video assessment scoring method to determine the accuracy of endoscopist performance at screening flexible sigmoidoscopy. Endoscopy 2006;38:218-225. doi:10.1055/s-2005-870445
22. Gross GWW, Bozymski EM, Brown RD, et al. Guidelines for training non-specialists in screening flexible sigmoidoscopy. Gastrointest Endosc 2000;51:783-785. doi:10.1053/ge.2000.v51.age516783
23. Doria-Rose VP, Newcomb PA, Levin TR. Incomplete screening flexible sigmoidoscopy associated with female sex, age, and increased risk of colorectal cancer. Gut 2005;54:1273-1278. doi:10.1136/gut.2005.064030
24. Ramakrishnan K, Scheid DC. Predictors of incomplete flexible sigmoidoscopy. J Am Board Fam Pract 2003;16:478-484. doi:10.3122/jabfm.16.6.478
25. Valliani T, Dowler K, Dunckley P. OC-021 Flexible sigmoidoscopy completion rates. Time to change certification criteria? Gut 2010;59(Suppl 1):A9. doi:10.1136/gut.2009.208934u
26. Barton JR, Corbett S, van der Vleuten CP; English Bowel Cancer Screening Programme; UK Joint Advisory Group for Gastrointestinal Endoscopy. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: assessing colonoscopic skills of senior endoscopists. Gastrointest Endosc 2012;75:591-597. doi:10.1016/j.gie.2011.09.053
27. Schoenfeld PS, Cash B, Kita J, Piorkowski M, Cruess D, Ransohoff D. Effectiveness and patient satisfaction with screening flexible sigmoidoscopy performed by registered nurses. Gastrointest Endosc 1999;49:158-162. doi:10.1016/S0016-5107(99)70480-3

Formative DOPS: Colonoscopy and Flexible Sigmoidoscopy

Date of procedure			
Trainee name		**Membership no. (eg. GMC/NMC)**	
Trainer name		**Membership no. (eg. GMC/NMC)**	
Outline of case			
Difficulty of case	**Easy**	**Moderate**	**Complicated**
Please tick appropriate box			

Level of supervision Complete DOPS form by ticking box to indicate the appropriate level of supervision required for each item below. Constructive feedback is key to this tool assisting in skill development.	**Maximal supervision** Supervisor undertakes the majority of the tasks/decisions & delivers constant verbal prompts	**Significant supervision** Trainee undertakes tasks requiring frequent supervisor input and verbal prompts	**Minimal supervision** Trainee undertakes tasks requiring occasional supervisor input and verbal prompts	**Competent for independent practice** no supervision required	**Not applicable**
Pre-procedure					
Indication					
Risk					
Confirms consent					
Preparation					
Equipment check					
Monitoring					
Sedation					
Comments					
Procedure					
Scope handling					
Tip control					
Air management					
Proactive problem solving					
Loop management					
Patient comfort					
Pace and progress					
Visualisation					
Comments					
Management of findings					
Recognition					
Management					
Complications					

For further information, please contact the JAG Office askjets@rcplondon.ac.uk 020 3075 1620 www.thejag.org.uk

Level of supervision	Maximal supervision	Significant supervision	Minimal supervision	Competent for independent practice	Not applicable
Post-procedure					
Report writing					
Management plan					
Comments					
ENTS (endoscopic non-technical skills)					
Communication and teamwork					
Situation awareness					
Leadership					
Judgement and decision making					
Comments					

Learning Objectives for the next case The objectives should be added to the trainee's personal development plan (PDP) once DOPS is completed	
1.	
2.	
3.	

Overall Degree of Supervision required	**Maximal Supervision** Supervisor undertakes the majority of the tasks/decisions & delivers constant verbal prompts	**Significant Supervision** Trainee undertakes tasks requiring frequent supervisor input and verbal prompts	**Minimal Supervision** Trainee undertakes tasks requiring occasional supervisor input and verbal prompts	**Competent for independent practice** no supervision required
Please tick appropriate box				

DOPS form descriptors

Pre Procedure	
Indication	• Assesses the appropriateness of the procedure and considers possible alternatives
Risk assessment	• Assesses co-morbidity including drug history • Assesses any procedure related risks relevant to patient • Takes appropriate action to minimise any risks
Confirms Consent	• Early in training the consent process should be witnessed by the trainer, once competent it is acceptable for the trainee to confirm that valid consent has been gained by another trained member of staff. • During the summative DOPS the process of obtaining consent should witnessed and assessed • Complete and full explanation of the procedure including proportionate risks and consequences without any significant omissions and individualised to the patient • Avoids the use of jargon • Does not raise any concerns unduly • Gives an opportunity for patient to ask questions by adopting appropriate verbal and non-verbal behaviours • Develops rapport with the patient • Respects the patient's own views, concerns and perceptions
Preparation	• Ensures all appropriate pre-procedure checks are performed as per local policies • Ensures that all assisting staff are fully appraised of the current case • Ensures that all medications and accessories likely to be required for this case are available
Equipment check	• Ensures the available scope is appropriate for the current patient and indication • Ensures the endoscope is functioning normally before attempting insertion
Monitoring	• Ensures appropriate monitoring of oxygen saturation and vital signs pre-procedure • Ensures appropriate action taken if readings are sub-optimal • Demonstrates awareness of clinical monitoring throughout procedure
Sedation	• When indicated inserts and secures IV access and uses appropriate topical anaesthesia • Uses sedation and/or analgesic doses in keeping with current guidelines and in the context of the physiology of the patient • Drug doses checked and confirmed with the assisting staff • Uses Nitrous Oxide (Entonox) appropriately*
Procedure	
Scope handling	• Exhibits good control of head and shaft of colonoscope at all times • Angulation controls manipulated using the left hand during the procedure • Demonstrates ability to use all scope functions (buttons/biopsy channel) whilst maintaining stable hold on colonoscope • Minimises external looping in shaft of instrument
Tip control	• **Integrated technique**: Combines tip and torque steering to accurately control the tip of colonoscope and manoeuvre the tip in the correct direction. • **Individual components:** • **Tip steering:** Avoids unnecessary mucosal contact and maintains luminal view, avoiding need for blind negotiation of flexures and 'slide-by' where possible • **Torque steering:** Demonstrates controlled torque steering using right hand/fingers to rotate shaft of colonoscope

	• **Luminal awareness:** Correctly identifies luminal direction using all available visual clues, and avoids red outs
Air management	• Appropriate insufflation and suction of air to minimise over-distension of bowel while maintaining adequate views
Pro-active problem solving	• **Anticipates** challenges and problems (e.g. flexures and loops) • Uses appropriate techniques and **strategies** to prevent problems and minimise difficulties and patient discomfort • **Recognition:** Early recognition of technical challenges and difficulties preventing progression (e.g. loops, fixed pelvis) • **Management:** Can articulate and demonstrate a logical approach to resolving technical challenges, including early change in strategy when progress not being made
Loop management	• Uses appropriate techniques (tip and torque steering, withdrawal, position change) to minimise and prevent loop formation • Early recognition of when loop is forming or has formed • Understands and can articulate techniques for resolution of loops • Resolves loops as soon as technically possible, to minimise patient discomfort and any compromise to scope function • Recognises when loop resolution not possible and safely inserts colonoscope with loop, with awareness and management of any associated patient discomfort
Pace and progress	• Takes sufficient time to maximise mucosal views • Insertion of colonoscope speed adjusted to minimise looping, prevent problems and manage difficulties • Able to complete both insertion and withdrawal at pace consistent with normal service lists, adjusted, depending on difficulty of procedure • Extent of examination is appropriate to the indication
Patient comfort	• Conscious awareness of patient discomfort and potential causes at all times • Applies logical strategy to minimise any potential or induced discomfort, including anticipation of problems and reducing patient anxiety • Able to utilise effective colonoscopy techniques to resolve the majority of pain-related problems without the need for increased analgesia • Appropriate escalation of analgesic use if technical strategies unsuccessful in managing patient discomfort
Visualisation	• Visually and digitally examines the rectum and perineum (or stomal) area to ensure no obstruction or contraindication to insertion of instrument • Well-judged and timely use of screen washes and water irrigation to ensure clear views • Utilises positional changes to maximise mucosal views • Ensures optimal luminal views throughout the examination • Uses mucosal washing and suction of fluid to ensure optimal visualisation of mucosa, particularly at potential blind spots (caecal pole, flexures, recto-sigmoid). • Retroversion in the rectum should be performed to fully visualise the lower rectum and dentate line. If rectal retroversion is not possible, the reason should be indicated. • Recognises and identifies landmarks of complete examination (appendix orifice, ileo-caecal valve, tri-radiate fold or anastomosis/neo-terminal ileum) • There is photo-documentation (or video) of significant findings and landmarks of completion

Management of Findings	
Pathology recognition	• Accurate determination of normal and abnormal findings • Appropriate use of mucosal enhancement techniques
Pathology management	• Takes appropriate specimens as indicated by the pathology and clinical context • Performs relevant therapy or interventions if appropriate in clinical context (includes taking no action) • For management of polyps please use DOPyS.
Complications	• Ensures risk of complications is minimised • Rapid recognition of complications both during and after the procedure • Manages any complications appropriately and safely
Post procedure	
Report writing	• Records a full and accurate description of procedure and findings • Extent of the procedure is recorded in the report and supported by image/video recording • Uses appropriate endoscopy scoring systems
Management plan	• Records an appropriate management plan (including medication, further investigation and responsibility for follow-up).
ENTS (endoscopic non-technical skills)	
Communication and teamwork	• Maintains clear communication with assisting staff • Gives and receives knowledge and information in a clear and timely fashion • Ensures that both the team and the endoscopist are working together, using the same core information and understand the 'big picture' of the case • Ensures that the patient is at the centre of the procedure, emphasising safety and comfort • Clear communication of results and management plan with patient and/or carers
Situation awareness	• Ensure procedure is carried out with full respect for privacy and dignity • Maintains continuous evaluation of the patient's condition • Ensures lack of distractions and maintains concentration, particularly during difficult situations • Intra-procedural changes to scope set-up monitored and rechecked
Leadership	• Provides emotional and cognitive support to team members by tailoring leadership and teaching style appropriately • Supports safety and quality by adhering to current protocols and codes of clinical practice • Adopts a calm and controlled demeanour when under pressure, utilising all resources to maintain control of the situation and taking responsibility for patient outcome
Judgement and decision making	• Considers options and possible courses of action to solve an issue or problem, including assessment of risk and benefit • Communicates decisions and actions to team members prior to implementation • Reviews outcomes of procedure or options for dealing with problems • Reflects on issues and institutes changes to improve practice