Accepted Manuscript

Submission Date: 2023-03-17
Accepted Date: 2023-10-26
Accepted Manuscript online: 2023-11-21

Endoscopy

Initiatives to increase colonoscopy capacity – is there an impact on polyp detection? A UK National Endoscopy Database analysis

Liya Lu, Jamie Catlow, Matthew D Rutter, Linda Sharp.

Affiliations below.

DOI: 10.1055/a-2214-9840

Please cite this article as: Lu L, Catlow J, Rutter M D et al. Initiatives to increase colonoscopy capacity – is there an impact on polyp detection? A UK National Endoscopy Database analysis. Endoscopy 2023. doi: 10.1055/a-2214-9840

Conflict of Interest: LS holds grants for endoscopy-related research from 3-D Matrix and Medtronic. The remaining authors declare that there is no conflict of interest declare

This study was supported by Health Foundation (http://dx.doi.org/10.13039/501100000724), AIMS ID 695428

Abstract:

Background

Mismatch between routine endoscopy capacity and demand means centres often implement initiatives to increase capacity, such as weekend working or using locums/agency staff (insourcing). There are concerns about whether increasing workload to meet demand could negatively impact quality. We investigated polyp detection, a key quality metric, in weekend vs weekday and insourced vs standard procedures using data from the UK National Endoscopy Database (NED).

Methods

We conducted a national retrospective cross-sectional study of diagnostic colonoscopies undertaken 01/01-04/04/2019. The primary outcome was mean number of polyps (MNP) and the secondary, polyp detection rate (PDR). Multi-level mixed-effect regression, fitting endoscopist as a random effect, was used to examine associations between procedure day (weekend/weekday) and type (insourced/standard) and these outcomes, adjusting for patient age, sex and indication.

Results

92,879 colonoscopies (weekends: 19,977 (21.5%); insourced: 9,909 (10.7%)) were performed by 2,496 endoscopists. For weekend colonoscopies, patients were more often female and less often having screening-related procedures; for insourced colonoscopies, patients were younger and less often attending for screening-related procedures (all $p<0.05$). Case-mix adjusted MNP was significantly lower for weekend vs weekday (IRR=0.86, (95%CI 0.83-0.89)) and for insourced vs standard procedures (IRR=0.91, (95%CI 0.87-0.95)). MNP was highest for weekday standard procedures and lowest for weekend insourced procedures, but there was no interaction between procedure day and type. Similar associations were found for PDR.

Conclusions

Strategies to increase colonoscopy capacity may have adverse effects on polyp detection. Routine quality monitoring should be undertaken following such initiatives. Meantime, reasons for this unwarranted variation require investigation.

Corresponding Author:

Prof. Linda Sharp, Newcastle University, Population Health Sciences Institute, Newcastle upon Tyne, United Kingdom of Great Britain and Northern Ireland, linda.sharp@newcastle.ac.uk

Affiliations:
Liya Lu, Newcastle University, Population Health Sciences Institute, Newcastle upon Tyne, United Kingdom of Great Britain and Northern Ireland
Jamie Catlow, Newcastle University, Population Health Sciences Institute, Newcastle upon Tyne, United Kingdom of Great Britain and Northern Ireland
Jamie Catlow, University Hospital of North Tees, Gastroenterology, Stockton-on-Tees, United Kingdom of Great Britain and Northern Ireland
[...]
Linda Sharp, Newcastle University, Population Health Sciences Institute, Newcastle upon Tyne, United Kingdom of Great Britain and Northern Ireland

Table 1. Patient characteristics overall and by whether procedures were undertaken at weekends, or insourced

	Total	Weekend			Insourced		
	(N=92,879) N(%)	Yes (N=19,977) N (%)	No (N=72,902) N (%)	P value for χ^2 test	Yes (N=9,909) N (%)	No (N=82,970) N (%)	P value for χ^2 test
Age, years				0.013			<0.001
18-39	11,649 (12.5)	2,367 (11.9)	9,282(12.7)		1,579 (15.9)	10,070 (12.1)	
40-49	11,540 (12.4)	2,522 (12.6)	9,018 (12.4)		1,525 (15.4)	10,015 (12.1)	
50-59	20,237 (21.8)	4,377 (21.9)	15,860 (21.7)		2,348 (23.7)	17,889 (21.5)	
60-69	23,461 (25.3)	5,031 (25.2)	18,430 (25.3)		2,166 (21.9)	21,295 (25.7)	
≥70	25,992 (28.0)	5,680 (28.4)	20,312 (27.9)		2,291 (23.1)	23,701 (28.6)	
Sex				0.001			0.072
Female[a]	46,861 (50.4)	10,283 (51.5)	36,578 (50.2)		5,084 (51.3)	41,777 (50.3)	
Male	46,018 (49.6)	9,694 (48.5)	36,324 (49.8)		4,825 (48.7)	41,193 (49.7)	
Indication				<0.001			<0.001
Screening	8,828 (9.5)	598 (3.0)	8,230 (11.3)		606 (6.1)	8,222 (9.9)	
IBD assessment	4,027 (4.3)	827 (4.4)	3,155 (4.3)		377 (3.8)	3,650 (4.4)	
Previous polyp(s)	9,760 (10.5)	2,079 (10.4)	7,681 (10.5)		909 (9.2)	8,851 (10.7)	
Previous abnormal investigation	2,551 (2.8)	570 (2.8)	1,981 (2.7)		276 (2.8)	2,275 (2.7)	
Lower GI symptoms	36,419 (39.2)	7,642 (38.3)	28,777 (39.5)		4,539 (45.8)	31,880 (38.4)	
Other[b]	31,294 (33.7)	8,216 (41.1)	23,078 (32.7)		3,202 (32.3)	28,092 (33.9)	

[a] included 1005 unknown sex

[b] polyposis, family history of colorectal cancer (CRC), CRC-follow up, stenting replacement or removal, tumour assessment, weight loss or other

Table 2. MNP and PDR, overall, and by whether procedures were undertaken at weekends, or insourced

	Total	Weekend		Insourced	
		Yes	No	Yes	No
MNP, Mean (95%CI)[a]	68.7 (68.0-69.5)	56.7 (55.2-58.2)	72.1 (0.71-0.73)	56.7 (54.6-58.8)	70.2 (69.3-71.0)
PDR, % (95%CI)[b]	34.9 (34.6-35.2)	31.1 (30.5-31.8)	35.9 (35.6-36.3)	30.9 (30.0-31.8)	35.4 (35.0-35.7)

[a] t-test for MNP: p<0.001

[b] χ^2 test for PDR: p<0.001

MNP, mean number of polyps per 100 colonoscopies; PDR, polyp detection rate

Table 3. Mixed effects regression analyses of associations between weekend/weekday and insourced/standard procedures and MNP and PDR: univariable and multivariable incidence rate ratios (IRR) and odds ratios (OR) with 95% confidence intervals

	MNP: primary outcome			PDR: secondary outcome		
	Univariable IRR (95%CI)	Multivariable IRR (95%CI)		Univariable OR (95%CI)	Multivariable OR (95%CI)	
		Case-mix adjusted[a]	Fully adjusted[b]		Case-mix adjusted[a]	Fully adjusted[b]
Patient level						
MNP, Mean (95%CI)	68.7 (68.0-69.5)		68.1 (67.7-68.5)			
PDR, % (95%CI)				34.9 (34.6-35.2)		34.7 (34.5-34.8)
Weekend colonoscopy						
No	1.00	1.00	1.00	1.00	1.00	1.00
Yes	0.81 (0.78-0.84)[c]	0.87 (0.84-0.90)[c]	0.86 (0.83-0.89)[c]	0.81 (0.78-0.85)[c]	0.86 (0.82-0.90)[c]	0.86 (0.82-0.90)[c]
Insourced colonoscopy						
No	1.00	1.00	1.00	1.00	1.00	1.00
Yes	0.81 (0.77-0.84)[c]	0.91 (0.87-0.95)[c]	0.91 (0.87-0.95)[c]	0.76 (0.72-0.80)[c]	0.90 (0.85-0.95)[c]	0.90 (0.85-0.95)[c]
Endoscopist level, variance (SE)			0.27 (0.01)			0.39 (0.02)

[a] adjusted for age, sex and indication only

[b] adjusted for age, sex and indication and, as appropriate, whether or not procedures were undertaken at weekend or insourced

[c] p<0.001

MNP, mean number of polyps per 100 procedures; PDR, polyp detection rate; IRR, incidence rate ratio; OR, odds ratio

Table 4. Comparison of other markers of colonoscopy quality and workload and by whether procedures were undertaken at weekends, or insourced

	Weekend			Insourced		
	Yes (N=19,977) N (%)	No (N=72,902) N (%)	P value	Yes (N=9,909) N (%)	No (N=82,970) N (%)	P value
CIR,%[a]	93.4	93.5	0.526[b]	94.7	93.4	<0.001[b]
TIR,%[a]	37.3	38.8	<0.001[b]	44.6	37.7	<0.001[b]
WT for negative procedures[c]**, minutes**			<0.001[b]			0.014[b]
<6	444 (3.3)	1,798 (3.9)		245 (3.7)	1,997 (3.8)	
6-9	8,205 (61.4)	21,357 (47.1)		3,298 (49.7)	26,264 (50.4)	
10-30	3,732 (28.0)	14,016 (30.9)		1,827 (27.5)	15,921 (30.6)	
Unknown	972 (7.3)	8,203 (18.1)		1,264 (19.1)	7,911 (15.2)	
Bowel preparation			<0.001[b]			<0.001[b]
Excellent	2,952 (14.8)	15,705 (21.6)		2,180 (22.0)	16,477 (19.9)	
Good	9,852 (49.3)	35,590 (48.8)		4,846 (48.9)	40,596 (48.9)	
Fair	5,834 (29.2)	17,398 (23.9)		2,364 (23.9)	20,868 (25.1)	
Inadequate	1,030 (5.2)	3,092 (4.2)		420 (4.2)	3,702 (4.5)	
Unknown	309 (1.5)	1,117 (1.5)		99 (1.0)	1,327 (1.6)	
Discomfort score			<0.001[b]			0.034[b]
None/minimal/mild	19,299 (96.6)	69,901 (95.9)		9,576 (96.6)	79,624 (96.0)	
Moderate/severe	512 (2.6)	2,603 (3.6)		297 (3.0)	2,818 (3.4)	
Unknown	166 (0.8)	398 (0.5)		36 (0.4)	528 (0.6)	
Buscopan use	2,903 (14.5)	14,670 (20.1)	<0.001[b]	1,989 (20.1)	15,584 (18.8)	0.002[b]
Intravenous sedation use	13,703 (68.6)	51,905 (71.2)	<0.001[b]	7,416 (74.8)	58,192 (70.1)	<0.001[b]
Procedure points on list, mean (SD)[d]	10.7 (3.3)	8.2 (3.4)	<0.001[e]	8.8 (3.5)	8.3 (3.8)	<0.001[e]

[a] denominator excluded 4 pouch procedures but otherwise was based on the full dataset without exclusions (N=99,996)

[b] χ^2 test

[c] based on 58,727 negative procedures.

[d] based on all 100,000 procedures

[e] t-test

CIR, caecal intubation rate; TIR, terminal ileum intubation rate; WT, withdrawal time

SUPPLEMENTARY MATERIAL

Liya Lu, Jamie Catlow, Matthew D Rutter, Linda Sharp, on behalf of the NED-APRIQOT study team

Initiatives to increase colonoscopy capacity – is there an impact on polyp detection? A UK National Endoscopy Database analysis

Supplementary Table 1. Secondary analysis: regression analyses[a] for association between the combination of weekend/weekday and insourced/standard colonoscopy and MNP and PDR

	MNP: primary outcome				PDR: secondary outcome			
	Univariable *Mean ± SD*	**Multivariable[b]** *Mean ± SD*	**Univariable** *IRR (95%CI)*	**Multivariable[b]** *IRR (95%CI)*	***Univariable*** *%*	***Multivariable[b]*** *%*	***Univariable*** *OR (95%CI)*	***Multivariable[b]*** *OR (95%CI)*
Patient level Colonoscopy								
Weekday standard	73.6±125.3	72.9±60.3	1.00	1.00	36.5	36.2	1.00	1.00
Weekday insourced	57.4±109.0	56.5±47.4	0.76 (0.72-0.80)[c]	0.90 (0.86-0.95)[c]	30.9	30.7	0.70 (0.66-0.75)[c]	0.88 (0.83-0.95)[c]
Weekend standard	57.0±108.3	56.8±44.4	0.79 (0.76-0.81)[c]	0.86 (0.83-0.89)[c]	31.2	31.0	0.77 (0.74-0.81)[c]	0.85 (0.81-0.90)[c]
Weekend insourced	55.0±104.7	54.9±40.7	0.73 (0.67-0.79)[c]	0.79 (0.73-0.87)[c]	30.7	30.6	0.71 (0.65-0.79)[c]	0.80 (0.72-0.88)[c]
Endoscopist level, variance (SE)			0.41 (0.02)	0.27 (0.01)			0.55 (0.02)	0.39 (0.02)

[a] multi-level models with procedure level variables as fixed effects and endoscopist as a random effect
[b] adjusted for case-mix (patient age, sex, indication)
[c] p<0.001
MNP, mean number of polyps; PDR, polyp detection rate; IRR, incidence rate ratio; OR, odds ratio

Supplementary table 2. Comparison of other markers of colonoscopy quality and workload for the combination of weekend/weekday and insourced/standard colonoscopy

	Weekday, standard (N=65,813) N (%)	Weekday, insourced (N=7,089) N (%)	Weekend, standard (N=17,157) N (%)	W
CIR[a]	66306 (93.3)	7115 (95.2)	17233 (93.4)	
TIR[a]	26943 (37.9)	3490 (46.7)	6818 (36.9)	
WT for negative procedures[c]**, minutes**				
<6	1613 (4.0)	185 (3.9)	384 (3.3)	
6-9	19239 (47.3)	2118 (44.7)	7025 (61.3)	
10-30	12754 (31.4)	1262 (26.7)	3167 (27.7)	
Unknown	7035 (17.3)	1168 (24.7)	876 (7.7)	
Bowel preparation				
Excellent	13881 (21.1)	1824 (25.7)	2596 (15.1)	
Good	32148 (48.8)	3442 (48.5)	8448 (49.3)	
Fair	15925 (24.2)	1473 (20.8)	4943 (28.8)	
Inadequate	2811 (4.3)	281 (4.0)	891 (5.2)	
Unknown	1048 (1.6)	69 (1.0)	279 (1.6)	
Discomfort score				
None/minimal/mild	63100 (95.9)	6801 (96.0)	16524 (96.3)	
Moderate/severe	2346 (3.5)	257 (3.6)	472 (2.8)	
Unknown	367 (0.6)	31 (0.4)	161 (0.9)	
Buscopan use	13078 (19.9)	1592 (22.5)	2506 (14.6)	
Intravenous sedation use	46423 (70.5)	5482 (77.3)	11769 (68.6)	
Procedure points on list, mean (SD)[d]	8.3 (3.4)	7.2 (3.4)	10.6 (3.3)	

[a] denominator excluded 4 pouch procedures but otherwise was based on the full dataset without exclusions (N=99996)
[b] χ^2 test
[c] based on 58,727 negative procedures.
[d] based on all 100000 procedures
[e] one-way ANOVA
CIR, caecal intubation rate; TIR, terminal ileum intubation rate; WT, withdrawal time

Initiatives to increase colonoscopy capacity – is there an impact on polyp detection? A UK National Endoscopy Database analysis

Liya Lu[1+], Jamie Catlow[1,2,3], Matthew D Rutter[1,2*], Linda Sharp[1*], on behalf of the NED-APRIQOT study team[4]

[1] Population Health Sciences Institute, Newcastle University, Newcastle-upon-Tyne, UK

[2] Gastroenterology, University Hospital of North Tees, Stockton-on-Tees, UK

[3] Gastroenterology, Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, UK

[4] NED-APRIQOT study team listed in Acknowledgements

[+] Present affiliation NHS Forth Valley, Stirling, UK

[*] Joint senior authors: L Sharp and MD Rutter

Running Title: Colonoscopy quality, insourcing and weekend working

Correspondence to: Linda Sharp, Room 5.15, Ridley 1 Building, Newcastle University, Claremont Road, Newcastle upon Tyne NE1 7RU; tel: +44 (0)191 208 6275; email: linda.sharp@ncl.ac.uk; ORCID: 0000-0001-9515-1722

Main manuscript word count: 3499

ABSTRACT

Background

Mismatch between routine endoscopy capacity and demand means centres often implement initiatives to increase capacity, such as weekend working or using locums/agency staff (insourcing). There are concerns about whether increasing workload to meet demand could negatively impact quality. We investigated polyp detection, a key quality metric, in weekend vs weekday and insourced vs standard procedures using data from the UK National Endoscopy Database (NED).

Methods

We conducted a national retrospective cross-sectional study of diagnostic colonoscopies undertaken 01/01-04/04/2019. The primary outcome was mean number of polyps (MNP) and the secondary, polyp detection rate (PDR). Multi-level mixed-effect regression, fitting endoscopist as a random effect, was used to examine associations between procedure day (weekend/weekday) and type (insourced/standard) and these outcomes, adjusting for patient age, sex and indication.

Results

92,879 colonoscopies (weekends: 19,977 (21.5%); insourced: 9,909 (10.7%)) were performed by 2,496 endoscopists. For weekend colonoscopies, patients were more often female and less often having screening-related procedures; for insourced colonoscopies, patients were younger and less often attending for screening-related procedures (all $p<0.05$). Case-mix adjusted MNP was significantly lower for weekend vs weekday (IRR=0.86, (95%CI 0.83-0.89)) and for insourced vs standard procedures (IRR=0.91, (95%CI 0.87-0.95)). MNP was highest for weekday standard procedures and lowest for weekend insourced procedures, but there was no interaction between procedure day and type. Similar associations were found for PDR.

Conclusions

Strategies to increase colonoscopy capacity may have adverse effects on polyp detection. Routine quality monitoring should be undertaken following such initiatives. Meantime, reasons for this unwarranted variation require investigation.

Keywords

colonoscopy; mean number of polyps detected; polyps detection rate; weekend; insourcing; quality

Introduction

Colonoscopy is the gold standard investigation for the large bowel. Its use has increased year on year. For example, in the UK, 700,000 procedures were conducted in the year before the COVID-19 pandemic [1] and demand is projected to continue growing [2].

This growth means endoscopy services are under pressure. There are variations in access across Europe [3] and, in the UK, services are failing to meet waiting time targets [4]. To increase capacity, many services have implemented special initiatives including: extending working hours on weekdays and at weekends (paying staff for overtime); subcontracting services to a supplier that employs locums or agency staff and uses the NHS organisation's premises and equipment to deliver procedures (insourcing); and referring patients to an external provider (outsourcing). Use of such initiatives is widespread; in 2021, 46% of UK services used insourcing and 53% conducted weekend lists [4-7].

Maximising colonoscopy quality is paramount to minimise missed cancers and maximise potential to prevent cancer through premalignant polyp detection and resection. Perhaps the most important colonoscopy performance measures relate to polyp detection: studies demonstrate an inverse correlation between adenoma detection rate (ADR) and post-colonoscopy colorectal cancer (PCCRC) rate [8,9]. Colonoscopy quality in Europe has improved [10-14], but concern has been raised about the potential impact on quality of increasing workload to address increased demand. Service pressures and financial incentives, particularly associated with waiting list initiatives or in/outsourcing services, might overtly or subconsciously favour "doing more" rather than "doing well" [5]. Quality might also reduce when an endoscopist practises outside their base unit, as happens with insourcing, due to relative unfamiliarity with equipment, staff and processes.

Elsewhere in clinical practice there are reports (albeit inconsistent) of a "weekend effect" characterised by lower standards of care and/or poorer outcomes for patients treated at weekends than weekdays (see, for example, [15-18]). A recent colonoscopy study reported that while quality standards were maintained for weekend and evening procedures, ADR and mean polyps per procedure were lower on Saturdays and evenings than on weekdays [19].

However, this single centre analysis did not adjust ADR for differences in characteristics of patients seen at weekends and weekdays (case-mix), which can induce artefactual differences in outcomes [20]. Investigations of quality and insourcing are lacking, but a review of research on use of locums in medical practice noted concerns about patient safety and higher risk of harms [21].

The UK's National Endoscopy Database (NED) is a novel registry that captures real-time patient-level data automatically from each hospital's endoscopy reporting system (ERS). NED, which commenced roll-out in 2016, aims to capture all endoscopic procedures from all (around 520 in both NHS and independent sector) UK endoscopy units [22]. Using NED data, we investigated whether insourcing and/or weekend working was associated with colonoscopy quality. Our aims were to: examine whether polyp detection rates vary between weekend and weekday procedures, and between regular versus insourcing activity, after accounting for case-mix; and investigate whether there was any interaction between procedure day (weekday/weekend) and type (standard/insourcing). In addition, we also aimed to explore possible associations between other pre-selected markers of colonoscopy quality and workload and procedure day and type.

Materials & Methods

Data source, design and inclusion criteria

NED collects data on endoscopic process such as indications, diagnoses, therapies and patient comfort in each procedure, and anonymised data on patient age and sex; data is recorded at procedure level; histology is not recorded [22].

For this retrospective cross-sectional study, within the NED-APRIQOT research project [23], we extracted data on 100,000 consecutive independent (i.e. non-training) colonoscopies between 01/01/2019-04/04/2019. This was the most recent quarter with available data when the study commenced. Data were available on the procedure and its outcome (e.g. whether conducted as a result of the Bowel Cancer Screening Programme (BCSP), urgency, therapeutics, polyps detected), the patient (e.g. age, sex, indication for colonoscopy), which endoscopist conducted the procedure and the unit/Trust in which it was done.

The dataset was restricted to: complete non-emergency procedures (i.e. colonoscopy that reported extent of caecum, terminal ileum, neo-terminal ileum or ileocolonic anastomosis), with withdrawal time <120 minutes and complete polypectomy data, conducted on patients aged 18-99 years (Supplementary Figure 1).

Outcome measures

We considered two recognised detection-related colonoscopy key performance indicators (KPIs): mean number of polyps detected (MNP, expressed per 100 procedures, primary outcome), and polyp detection rate (PDR, secondary outcome). MNP was defined as total number of polyps detected divided by total number of colonoscopies multiplied by 100. As a metric, it aligns more closely to what is being assessed (i.e. thorough inspection of the entire colorectum) than "one and done" metrics such as ADR [23]. A cap of five polyps per procedure was applied to ensure comparisons were not unduly influenced by polyposis patients and in line with evidence that ADR plateaus around five polyps detected [24]. PDR, which is considered an acceptable surrogate for ADR in the absence of histology [25], was defined as percentage of colonoscopies in which at least one polyp was detected. Examples of polyps are shown in Figures 1 and 2.

Main explanatory variables

Each colonoscopy was classified according to whether it was (i) performed on a weekend (variable: weekend Yes (performed Saturday/Sunday)/No (performed Monday-Friday)) and (ii) an insourced or standard procedure (variable: insourced Yes/No). For each endoscopist, the hospital Trust where they performed most procedures during the analysis time period, was considered their main Trust. If an endoscopist worked in only one Trust, that Trust was their main Trust. Standard colonoscopies were defined as colonoscopies conducted in an endoscopist's main Trust. Insourced colonoscopies were defined as colonoscopies conducted in an endoscopist's non-main Trust.

Other endoscopic data

Polyps are more common in men and older individuals, and prevalence varies by indication [26]. Adjusting outcomes for case-mix is therefore important. Three case-mix variables were considered: patient sex, age and procedure indication. Procedures where sex was unknown were combined with females for analysis. Patient age was categorised as: 18-39, 40-49, 50-59, 60- 69, and 70+ years. A hierarchical indication variable was derived: screening (indication BCSP or faecal occult blood test (FOB); the study was conducted before widespread use of faecal tests for triage of symptomatic populations); inflammatory bowel disease (IBD) assessment (IBD assessment/ surveillance only or plus any other indications recorded but not screening); previous polyps (previous polyps only or plus other indications but neither screening nor IBD); previous abnormal investigation (previous abnormal sigmoidoscopy or abnormal CT only or plus other indications but not screening, IBD or previous polyps); lower gastrointestinal (GI) symptoms (constipation, diarrhoea, chronic alternating diarrhoea/constipation, previous bleeding, abdominal pain, abdominal mass, or anaemia only or plus other indications but not screening, IBD, previous polyps, or previous abnormal investigations); and other (polyposis, family history of colorectal cancer (CRC), CRC-follow up, stenting replacement or removal, tumour assessment, weight loss or other).

Other markers of colonoscopy quality and workload [25] were considered for purposes of interpretation. Each was computed separately for weekend (yes/no) and insourced (yes/no) procedures. Caecal intubation rate (CIR) and terminal ileum intubation rate (TIR) were the percentage of colonoscopies in which the colonoscope tip passed at least to the caecum and neo-terminal ileum or terminal ileum, respectively. Withdrawal time was based on negative procedures only (i.e. no polyps were found and no therapeutics were done) and grouped into <6, 6 to 9, 10-30 minutes and unknown. Bowel preparation quality score was categorised as excellent, fair, good, inadequate and unknown. Patient discomfort score was assessed by an endoscopist or a nurse or both (in which instance, the worst score was used for analysis); it was grouped as none/minimal/mild, moderate/severe, and unknown. Hyoscine butylbromide (Buscopan) and sedation (pethidine, midazolam or fentanyl) use were also summarised.

Procedure points were used to describe list length with two points allocated for the index colonoscopy; one additional point allocated for each other upper endoscopy or flexible

sigmoidoscopy on the list; two additional points for each additional colonoscopy; and three points for each ERCP; list length was capped at 16 points, as the longest plausible length.

Ethical Approval

Health Research Authority and Health and Care Research Wales approval was obtained [23].

Statistical analysis

Descriptive statistics were calculated for the study population for case-mix, other quality and workload variables. For each variable, values were compared by procedure day (weekend/weekday) or type (insourced/standard) using chi-square tests or t-tests as appropriate.

Given the hierarchical structure of the dataset, two-level mixed-effect regression models with procedure-level covariates as fixed effects and endoscopist as a random effect were used [27]. MNP is count data, but was over-dispersed (variance exceeded mean), so negative binomial regression was used. PDR is binary, therefore logistic regression was used. To investigate whether insourcing was associated with each outcome, we ran an unadjusted regression model, then adjusted for case-mix. We repeated this for the weekend variable. The interaction between weekend/weekday and insourcing/standard was tested by fitting an interaction term; this was not statistically significant, so the final models include only main effects. The final case-mix adjusted weekend/weekday model was also adjusted for insourcing/standard (and vice versa). For information, we ran a secondary analysis fitting a 2x2 variable for the combinations of weekend/weekday and insourcing/standard, adjusted for case-mix. Finally, to consider the possibility that the results may be affected by potential confounders we could not control for (such as family history or presence of familial syndromes), we conducted a sensitivity analysis where we re-calculated the main effects of procedure day and type, restricting consideration to procedures undertaken in patients aged 50 and older. Throughout, $p \leq 0.05$ (two-sided) was considered statistically significant; no adjustments were made for multiple testing. All statistical analyses were performed using Stata 16.1 (StataCorp, Texas, USA).

Results

Procedure Characteristics

After exclusions, the analysis included 92,879 colonoscopies performed by 2,496 endoscopists from 332 units and 112 Trusts (Supplementary Figure 1). Fifty-three percent were conducted among patients ≥60 years and half were in male patients. Investigation of lower GI symptoms was the indication for 39.2%; surveillance of previous polyps accounted for 10.5% and screening-related colonoscopies for 9.5% (Table 1).

Just over one-fifth (21.5%) of colonoscopies were performed at weekends. Just over one in ten were insourced (10.7%). Overall, 70.9% (65,813) were standard and done on weekdays, 18.5% (17,157) were standard and on weekends, 7.6% were weekday insourced (7,089), and 3.0% (2,820) were weekend insourced.

In weekend colonoscopies, patients were less often male or attending for a screening-related procedure (Table 1). Patient age differed little between weekend and weekdays. In insourced colonoscopies, patients were younger and less likely to be attending for a screening-related procedure.

Primary outcome: weekends, insourcing and MNP

The overall MNP was 68.7 per 100 procedures (Table 2). MNP was lower in weekend than weekday colonoscopies (57 vs 72 per 100 procedures), and in insourced than standard colonoscopies (57 vs 70 per 100 procedures).

In multivariable case-mix adjusted analyses, MNP was statistically significantly lower at weekends compared with weekdays (Table 3). This association persisted after further adjustment for insourcing (incidence rate ratio [IRR]=0.86, (95%CI 0.83-0.89)). Likewise, there was a statistically significant association between insourced colonoscopies and lower MNP for weekend procedures (IRR=0.91, (95%CI 0.87-0.95)), suggesting weekend and insourcing have independent deleterious effects on MNP.

Older patient age, the patient being male, and procedure indications of screening, previous polyps and previous abnormal investigation were significantly associated with higher MNP (data not shown).

Secondary outcome: weekends, insourcing and PDR

The overall PDR was 34.9% (Table 2). It was lower at weekends and for insourced procedures (weekends 31.1%; weekdays 35.9%; insourced 30.9%; standard 35.4%).

The final case-mix adjusted multivariable models showed significantly lower PDR in both weekend procedures (OR=0.86, (95%CI 0.82-0.90)) and insourced procedures (OR=0.90, (95%CI 0.85-0.95))(Table 3).

Sensitivity analysis

When restricting consideration to procedures in people aged 50+ (n=69,690), effect estimates were little changed (fully adjusted - weekend vs weekday: MNP, IRR=0.86, (95%CI 0.83-0.90); PDR, OR=0.87, (95%CI 0.82-0.91); insourced vs standard: MNP, IRR=0.91, (95%CI 0.87-0.96); PDR, OR=0.92, (95%CI 0.86-0.98)).

Secondary analysis: combination of weekend and insourcing

Supplementary Table 1 shows the secondary analyses, where procedures were simultaneously categorised by whether they were weekend/weekday or insourced/standard. There was no statistically significant interaction between procedure day and type.

Other quality & workload indicators

Table 4 compares other markers of colonoscopy quality and workload by whether procedures were weekend/weekday or insourced/standard. TIR, but not CIR, was lower at weekends than weekdays. Two-thirds of negative weekend procedures had a withdrawal time <10 minutes versus half of weekday procedures. At weekends, a lower proportion of procedures had excellent/good bowel preparation and hyoscine butylbromide and sedation were less often used. Mean number of points per list was higher at weekends (10.7(±3.3) vs 8.2(±3.4)).

CIR and TIR were both higher for insourced than standard procedures. Slightly more insourced procedures had a withdrawal time <10 minutes. A higher proportion of insourced procedures had excellent or good bowel preparation, and both hyoscine butylbromide and sedation were

more common for insourced procedures, while fewer patients had moderate/severe discomfort. The mean number of points per list for insourced procedures was slightly higher (8.8(±3.5) vs 8.3(±3.8)).

Supplementary table 2 shows quality and workload indicators by combinations of weekend and insourcing.

Discussion

To address increasing endoscopy demand, weekend and insourced working are increasingly used. However, increasing capacity must not result in poorer service quality. To our knowledge, this is the first study to investigate whether colonoscopy quality, measured in terms of detection, is maintained in both weekend and insourced working. Our nationwide analysis included almost 100,000 procedures from almost 2,500 endoscopists at more than 300 endoscopy units. Crucially, we adjusted for case-mix, which is not under the control of the endoscopist or unit and, as we have shown here, varies between weekend and weekday, and insourced and standard, procedures.

The statistically significantly lower MNP and PDR for colonoscopies at weekends suggest a "weekend effect" on polyp detection. Independent of this effect, there was a statistically significant association between insourced working and lower MNP and PDR. These findings suggest that strategies being used to increase colonoscopy capacity are having adverse effects on polyp detection.

Although concerns have been raised about the potential quality y impact of strategies to increase workload, relatively few studies have examined this. While insourcing is used in almost half of UK units [4], no other studies appear to have examined quality of insourced colonoscopy procedures. Regarding weekend working, a single-centre UK study of 17,634 patients who underwent non-screening colonoscopies January 2016-November 2018, reported higher MNP and ADR during weekdays (0.49, 28.8%, respectively) than evenings (0.38, 24.2%) and Saturdays (0.39, 24.4%) in univariate analyses, but these differences were not statistically significant after adjusting for working team and bowel preparation [19]. A single-centre Chinese study of 34,022 screening colonoscopies compared ADR across weekdays [28]; taking Monday as the reference,

in an analysis adjusted for case-mix, bowel preparation and sedation, ADR was statistically significantly lower at weekends (OR=0.78, (95% CI 0.63-0.95)).

We examined other quality markers to shed light on possible explanations for the observed unwarranted variation in quality, and to inform what might be done to address this. We did not adjust for these factors because we did not wish to convey an impression that the variation observed can be "explained" (albeit statistically). Regarding procedure day, several quality markers (including withdrawal time for negative procedures, and use of hyoscine butylbromide) – were worse for weekend than weekday procedures. Moreover, mean number of points per list was also higher at weekends, consistent with other studies which indicate that greater (cumulative) workload may be associated with endoscopist fatigue, or pressure to save time by cutting corners, and hence lead to lower detection [28-30]. Many of these markers have been associated with ADR in UK endoscopy practice [31] and, unlike case-mix, they are, at least to some extent, under the control of the endoscopist or unit and, hence, potentially modifiable. In *post hoc* analyses we added several of these markers to the fully-adjusted models (bowel preparation, discomfort score, Buscopan use, intravenous sedation use and points on list). The main effects for weekend were slightly attenuated but remained statistically significant (MNP: IRR=0.92 (95%CI 0.89-0.96); PDR: OR=0.93 (95%CI 0.89-0.97)), suggesting these variables do not fully explain the observed weekend effects. Additional explanations should be considered.

Notably, similar patterns were not seen for insourced procedures; in some instances these other quality indicators were statistically significantly *better* for insourced procedures. Reflecting this, in *post hoc* analyses adjusted for other markers of quality and workload, risk estimates for insourced procedures were slightly further from unity (not shown). This suggests that rather than inherently lower quality of insourcing endoscopists *per se*, the reduced detection more likely reflects other factors such as the environment, list processes, or subtle differences in endoscopist mind-set or behaviours during insourcing lists (e.g. focusing on getting procedures done to complete the list and get home).

Implications

Heightened attention should be paid to monitoring quality in relation to weekend and insourced working - and for any other future initiatives to address endoscopy capacity issues, or backlogs. Moreover, the possibility of perverse (albeit unintended) incentives to cut corners or over-populate lists should always be considered in relation to capacity initiatives.

As regards improving withdrawal time and hyoscine butylbromide use, reiterating quality targets is unlikely to effect change; information alone is generally insufficient to change behaviour [32]. Instead, active interventions will be required. The QIC study found that introducing a low-cost "bundle" of evidence-based measures into routine colonoscopy practice increased hyoscine butylbromide use and ADR with the effects sustained over time particularly in endoscopists who had poorest performance pre-intervention [33]. Endoscopist feedback results in modest performance improvements [34]. The NED-APRIQOT study is testing an automated, tailored, audit and feedback-based intervention developed using empirical evidence on what influences colonoscopy behaviour and psychological theories [23, 35]; results are awaited.

Potential solutions for lower quality in insourced procedures are less obvious. There is an urgent need to further investigate what underlies the observed differences. Endoscopist experience, workload and working environment (e.g. working in a unfamiliar team and physical environment) require investigation. Here, qualitative research could illuminate attitudinal, psychosocial or experiential issues which influence behaviours and detection.

Limitations

NED does not contain histology so we could not calculate ADR. MNP is strongly associated with ADR [24] while low PDR has been associated with increased PPCRC [36]. Moreover, lower detection at weekends and in insourced colonoscopies was seen for both MNP and PDR. However, findings may have differed had histology been available; large-scale studies examining effects of initiatives to increase colonoscopy capacity on MAP and ADR are warranted.

Although NED is a nationwide database, not all Trusts were uploading during the study period.. However, there is no reason to believe associations between procedure day and type and

detection would vary between Trusts which were, and were not, uploading. Observed associations between procedure day and patient outcomes using routine administrative data may be an artefact of differences in quality of data recorded for weekdays and weekends [20]. NED data is uploaded in real-time by direct feed from the local electronic ERS, completed by the endoscopist; this is likely a consistently reliable record of procedural findings. However, insourcing endoscopists could be less familiar with the local ERS, increasing risk of data entry error. Weekend/weekday procedures were identified based on procedure date entered into the ERS; these could be subject to some random error but differential misclassification seems unlikely.

Just over 1% of procedures did not have patient sex recorded. *A priori* we decided to group these with females, rather than introduce bias by dropping them. When we re-ran the analysis grouping them with males, results were unchanged (not shown).

Because the study population was large, care should be taken to not assume that statistical significance equates with clinical significance. Some statistically significant differences in the other quality marker by weekend/weekday and insourced/standard were modest. However, we would consider the magnitude of differences in detection (>4% in PDR; >10 per 100 procedures for MNP) clinically meaningful and - given the association between PDR and PCCRC [36] - likely to have implications for patient outcomes.

Our dataset comprised a combination of screening-related and other procedures. In the UK, the former have a different case-mix, are highly quality-assured and result in a higher ADR. We did not stratify by screening because the numbers of weekend and insourced procedures were relatively small in the screening population. In *post hoc* analyses, limited to procedures undertaken within the BCSPs, effects largely persisted (fully adjusted model - weekend vs weekday: MNP, IRR=0.80, (95%CI 0.71-0.90); PDR, OR=0.69, (95%CI 0.55-0.86); insourced vs standard: MNP, IRR=0.81, (95%CI 0.72-0.92); PDR, OR=0.86, (95%CI 0.68-1.08)). However, these findings are based on only 8,828 procedures so should be considered tentative and further analyses, in larger series, are warranted.

Conclusions

This national analysis found evidence of lower quality, as measured by polyp detection, in both weekend and insourced colonoscopies, after case-mix adjustment. The findings demonstrate the importance of considering the potential unintended consequences of any initiatives to extend colonoscopy capacity and of routine monitoring of quality when such initiatives are implemented. Active interventions to promote good practice, and research to better understand what underlies unwarranted variations in quality in weekend and insourced procedures, are also required.

Acknowledgements

The authors would like to acknowledge the NED committee, JAG and all UK endoscopy sites that upload data to the National Endoscopy Database.

The NED-APRIQOT study team comprises, in addition to the authors: Dr Rashmi Bhardwaj-Gosling; Professor Matthew Brookes; Mr Raphael Broughton; Ms Jill Deane; Dr Paul Dunckley; Dr Joanne Gray; Ms Anna Jenkins; Dr Adetayo Kasim; Dr Tom Lee; Dr Stephen McCarthy; Dr Richard McNally; Ms Kay Milburn; Professor Eva Morris; Dr Emmanuel Ogundimu; Mr Peter Rogers; Dr Josephine Ruwende; Mr Simon Sinclair; Professor Falko Sniehotta; Dr Roland Valori; Dr Janelle Wagnild

Author contributions

Conceptualisation: MDR & LS; Funding: MDR & LS; Methodology: all authors; Formal analysis: LL; Writing original draft: LL; Review and Editing: JC, MDR, LS. All authors approved the final version for submission.

Funding

This study was funded by a grant from the Health Foundation for the NED-APRIQOT study (AIMS ID 695428). The funder has no role in the design of the study, the analysis, interpretation or the decision to submit for publication.

Data Availability Statement

The data was provided to the authors for use in the NED-APRIQOT study and they are unable to share it. However, data extracts from NED are available upon reasonable request to the Joint Advisory Group for Gastrointestinal Endoscopy (JAG).

Conflict of Interests

LS holds grants for endoscopy-related research from 3-D Matrix and Medtronic. The remaining authors declare that there is no conflict of interest. declare.

References

1. Beaton D, Sharp L, Trudgill NG, et al. UK endoscopy workload and workforce patterns: is there potential to increase capacity? A BSG analysis of the National Endoscopy Database. Frontline Gastroenterol [Epub ahead of print] doi: 10.1136/flgastro-2022-102294
2. Richards M. Diagnostics: Recovery and Renewal – Report of the Independent Review of Diagnostic Services for NHS England. NHS England, 2020. Available at URL: https://www.england.nhs.uk/publication/diagnostics-recovery-and-renewal-report-of-the-independent-review-of-diagnostic-services-for-nhs-england/ [Accessed 17/3/2022]
3. Norgine. Scope for change: what next for European colonoscopy services. 2017. Available at URL: https://norgine.com/wp-content/uploads/2017/09/Scope-for-change-what-next-for-European-colonoscopy-services.pdf [Accessed 17/03/2023]
4. Ravindran S, Thomas-Gibson S, Bano M, et al. The national census of UK endoscopy services 2021 Frontline Gastroenterol 2022;13:463–470
5. Shenbagaraj L, Thomas-Gibson S, Stebbing J, Broughton R, Dron M, Johnston D, Shaw T, Nadim H, Green JT. Endoscopy in 2017: a national survey of practice in the UK. Frontline Gastroenterol 2019;10:7-15
6. Ravindran S, Bassett P, Shaw T, et al. National census of UK endoscopy services in 2019. Frontline Gastroenterol 2021;12:451–60
7. Oats B. Getting it Right First Time (GIRFT) GIRFT Programme National Specialty Report Available at URL: https://gettingitrightfirsttime.co.uk/medical_specialties/gastroenterology/ [Accessed 17/3/2022]
8. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer N Engl J Med 2010;362:1795-803.
9. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and the risk of colorectal cancer and death. N Engl J Med 2014; 370:1298-306.
10. Spada C, Koulaouzidis A, Hassan C, et al. Colonoscopy quality across Europe: a report of the European Colonoscopy Quality Investigation (ECQI) Group. Endosc Int Open 2021; 9(10): E1456-E1462.
11. Kaminski M F, Thomas-Gibson S, Bugajski M et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2017; 49: 378–397.
12. Valori R, Cortas G, de Lange T et al. Performance measures for endoscopy services: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy. 2018; 50: 1186–1204.
13. Siau K, Green JT, Hawkes ND, Broughton R, et al., on behalf of the Joint Advisory Group on Gastrointestinal Endoscopy. Impact of the Joint Advisory Group on Gastrointestinal

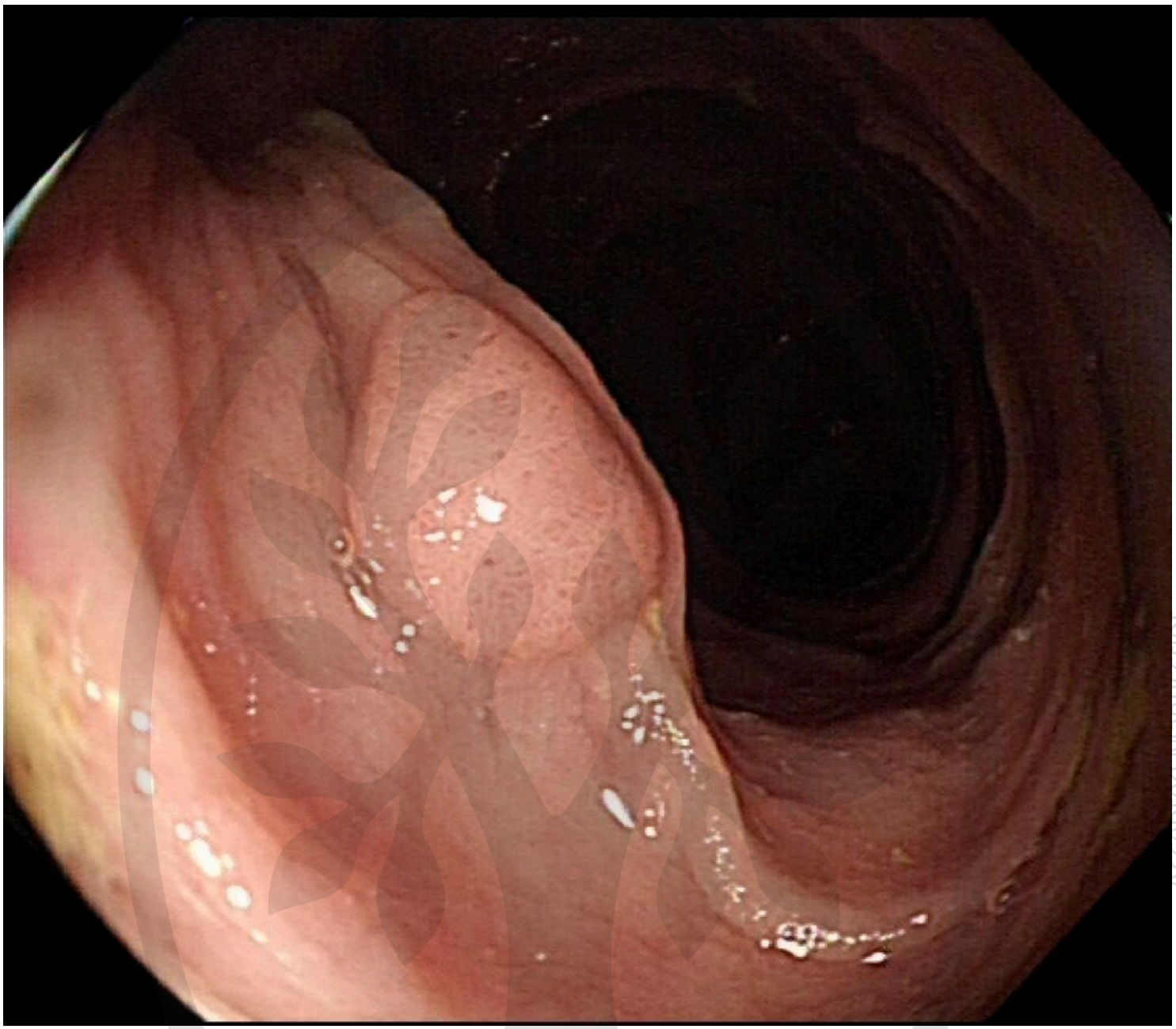
Endoscopy (JAG) on endoscopy services in the UK and beyond. Frontline Gastroenterol 2019;10:93-106
14. Burr NE, Derbyshire E, Taylor J, et al. Variation in post-colonoscopy colorectal cancer across colonoscopy providers in English National Health Service: population-based cohort study. BMJ. 2019;367:l6090
15. Lapointe-Shaw L, Abushomar H, Chen X-K, et al. Care and outcomes of patients with cancer admitted to the hospital on weekends and holidays: A retrospective cohort study. J Natl Compr Canc Netw 2016;14:867-74
16. Takagi H, Ando T, Umemoto T, ALICE (All-Literature Investigation of Cardiovascular Evidence) group. A meta-analysis of weekend admission and surgery for aortic rupture and dissection Vasc Med 2017;22:398-405
17. Chen Y-C, Armoiry X, Higenbottam C, et al. Magnitude and modifiers of the weekend effect in hospital admissions: a systematic review and meta-analysis BMJ Open 2019;9(6):e025764
18. Farrow L, Hall A, Aucott L, Holt G, Myint PK. Does quality of care in hip fracture vary by day of admission? Arch Osteoporos 2020;15:52
19. Afify SA, Abo-Elazm OM, Bahbah II, Thoufeeq MO. Weekend and evening planned colonoscopy activity: a safe and effective way to meet demands Endosc Int Open 2021;9:E1026–E1031
20. Li L, Rothwell PM, Study OV. Biases in detection of apparent "weekend effect" on outcome with administrative coding data: population-based study of stroke. BMJ 2016;353:i2648
21. Ferguson J, Walshe K. The quality and safety of locum doctors: a narrative review. J R Soc Med 2019;112:462–471
22. Lee TJ, Siau K, Esmaily S, et al. Development of a national automated endoscopy database: the United Kingdom national endoscopy database (NED). United European Gastroenterol J 2019;7:798–806
23. Catlow J, Sharp L, Kasim A, et al. The National Endoscopy Database (NED) Automated Performance Reports to Improve Quality Outcomes Trial (APRIQOT) randomized controlled trial design. Endosc Int Open 2020;8:E1545-E1552
24. Amano T, Nishida T, Shimakoshi H, Shimoda A, Osugi N, Sugimoto A, et al. Number of polyps detected is a useful indicator of quality of clinical colonoscopy. Endosc Int Open 2018;6:E878–84
25. Rees CJ, Thomas Gibson S, Rutter MD, et al., and British Society of Gastroenterology, the Joint Advisory Group on GI Endoscopy, the Association of Coloproctology of Great Britain and Ireland UK key performance indicators and quality assurance standards for colonoscopy. Gut 2016;65:1923-1929
26. Lee TJW, Rutter MD, Blanks RG, Moss SM, Goddard AF, Chilton A, et al. Colonoscopy quality measures: Experience from the NHS Bowel Cancer Screening Programme. Gut 2012;61:1050–7

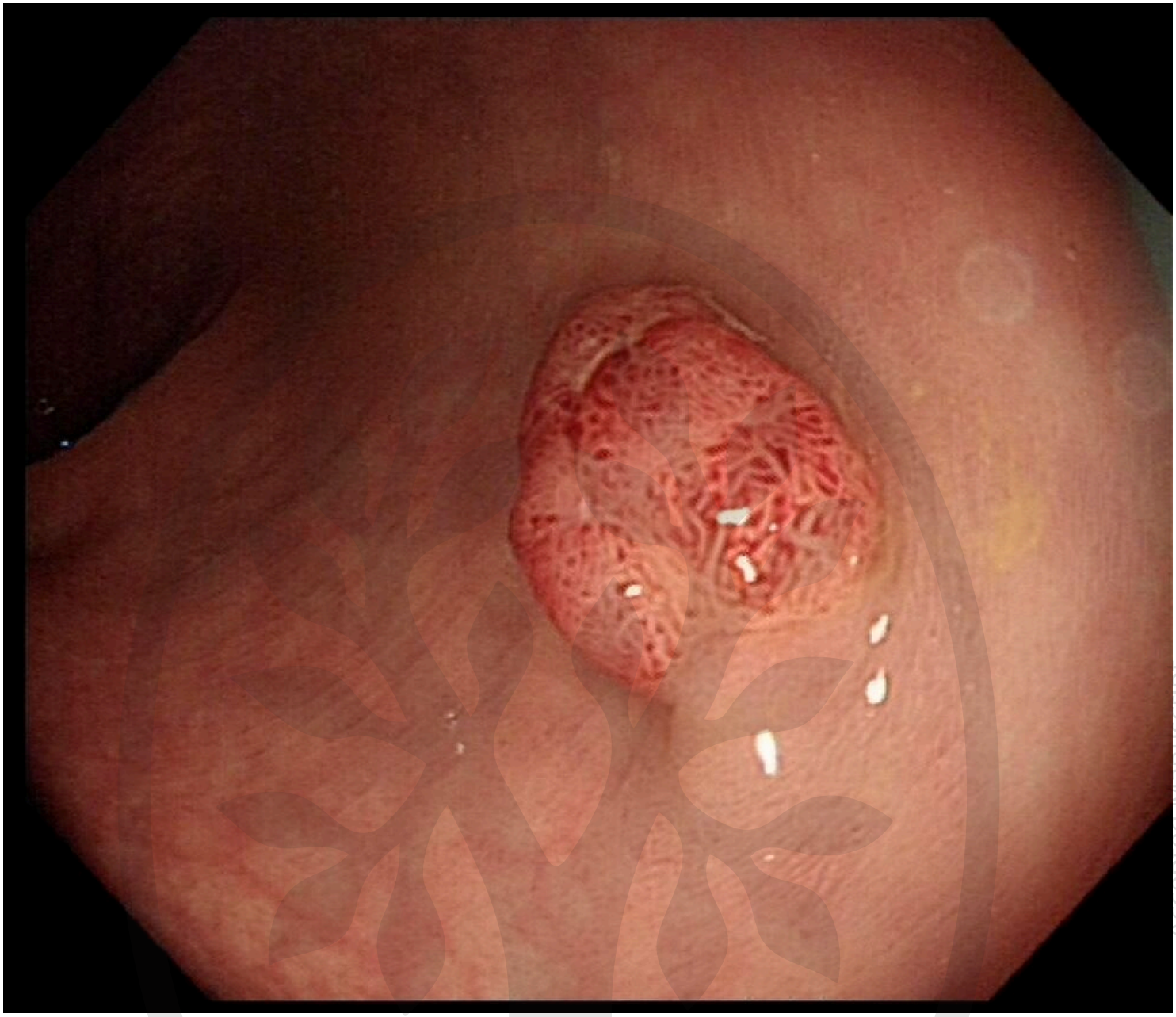
27. Brown H, Prescott R. Applied Mixed Models in Medicine, in Statistics in Practice. 2006, John Wiley & Sons: West Sussex UK
28. Dong Z, Wang J, Chen Y, et al. Negative effects of endoscopists' fatigue on colonoscopy quality on 34,022 screening colonoscopies. J Gastrointestin Liver Dis 2021;30:358-65
29. Teng TY, Khor SN, Kailasam M, Cheah WK, Lau CCC. Morning colonoscopies are associated with improved adenoma detection rates. Surg Endosc 2016;30:1796-803
30. Marcondes FO, Gourevitch RA, Schoen RE, Crockett SD, Morris M, Mehrotra A. Adenoma detection rate falls at the end of the day in a large multi-site sample. Dig Dis Sci 2018;63:856-9
31. Lee TJW, Rees CJ, Blanks RG, et al. Colonoscopic factors associated with adenoma detection in a national colorectal cancer screening program. Endoscopy 2014;46(3):203-11
32. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016. 136:109-16
33. Neilson LJ, East JE, Rajasekhar PT, et al. Sustained colonoscopy quality improvement using a simple intervention bundle. Endoscopy 2020;52(4):285-292
34. Bishay K, Causada-Calo N, Scaffidi MA, et al. Associations between endoscopist feedback and improvements in colonoscopy quality indicators: a systematic review and meta-analysis. Gastrointest Endosc 2020;92(5):1030-1040.e
35. Catlow J, Bhardwaj-Gosling R, Sharp L, Rutter MD, Sniehotta FF. Using a dark logic model to explore adverse effects in audit and feedback: a qualitative study of gaming in colonoscopy. BMJ Qual Saf 2022;31:704-15
36. Schwarz S, Hornschuch M, Pox C, Haug U. Polyp detection rate and cumulative incidence of post-colonoscopy colorectal cancer in Germany. Int J Cancer 2022;152:1547-1555.

Figure Legends

Figure 1. Paris 1s polyp.

Figure 2. 10mm Paris 1pm adenoma.





Supplementary Figure 1. Flowchart of study population