

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



Authors

Liya Lu^{1,*}, Jamie Catlow^{1,2,3}, Matthew D. Rutter^{1,2,‡}, Linda Sharp^{1,†}, on behalf of the NED-APRIQOT study team[†]

Institutions

- 1 Population Health Sciences Institute, Newcastle University, Newcastle-upon-Tyne, United Kingdom
- 2 Gastroenterology, University Hospital of North Tees, Stockton-on-Tees, United Kingdom
- 3 Gastroenterology, Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, United Kingdom

submitted 17.3.2023

accepted after revision 26.10.2023

accepted manuscript online 21.11.2023

published online 2024

Bibliography

Endoscopy

DOI 10.1055/a-2214-9840

ISSN 0013-726X

© 2024. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Supplementary Material

Supplementary material is available under
<https://doi.org/10.1055/a-2214-9840>

Corresponding author

Linda Sharp, PhD, Population Health Sciences Institute,
Newcastle University, Room 5.15, Ridley 1 Building,
Claremont Road, Newcastle-upon-Tyne NE1 7RU, United
Kingdom
linda.sharp@ncl.ac.uk

ABSTRACT

Background To address mismatch between routine endoscopy capacity and demand, centers often implement initiatives to increase capacity, such as weekend working or using locums/agency staff (insourcing). However, there are concerns that such initiatives may negatively impact quality. We investigated polyp detection for weekend vs. weekday and insourced vs. standard procedures using data from the UK National Endoscopy Database.

Methods We conducted a national, retrospective, cross-sectional study of diagnostic colonoscopies performed during 01/01–04/04/2019. The primary outcome was mean number of polyps (MNP) and the secondary outcome was 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 less often male or undergoing screening-related procedures; for insourced colonoscopies, patients were younger and less often undergoing screening-related procedures (all $P < 0.05$). Fully adjusted MNP was significantly lower for weekend vs. weekday (incidence rate ratio [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; there was no interaction between procedure day and type. Similar associations were found for PDR.

Conclusions Strategies to increase colonoscopy capacity may negatively impact polyp detection and should be monitored for quality. Reasons for this unwarranted variation require investigation.

* Present affiliation NHS Forth Valley, Stirling, United Kingdom

‡ Linda Sharp and Matthew D. Rutter are joint senior authors.

† NED-APRIQOT study team listed in Acknowledgments

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 organization'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].

Maximizing colonoscopy quality is paramount in minimizing the number of missed cancers and maximizing the 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 rate [8,9]. Colonoscopy quality in Europe has improved [10–14], but concern has been raised about the potential impact on quality of initiatives that are introduced to increase the workload to address the increase in demand. Service pressures and financial incentives, particularly associated with waiting list initiatives or in-/outsourcing services, might overtly or subconsciously favor “doing more” rather than “doing well” [5]. Quality might also reduce when an endoscopist practices 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” characterized by lower standards of care and/or poorer outcomes for patients treated at weekends compared with patients treated during the week (see, for example [15–18]). A recent colonoscopy study reported that while quality standards were maintained for weekend and evening procedures, ADR and mean number of polyps per procedure were lower on Saturdays and evenings than on weekdays [19]. However, this single-center analysis did not adjust ADR for differences in characteristics of patients seen at weekends and weekdays (case-mix), which can induce artifactual 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. NED, which commenced roll-out in 2016, aims to capture all endoscopy procedures from all UK endoscopy units (about 520 across both the state National Health Service [NHS] and the indepen-

dent sector) [22]. Using NED data, we investigated whether insourcing and/or weekend working was associated with colonoscopy quality. Our aims were to: 1) examine whether polyp detection rates vary between weekend and weekday procedures, and between regular and insourcing activity, after accounting for case-mix; and 2) 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, procedure day, and procedure type.

Methods

Data source, design, and inclusion criteria

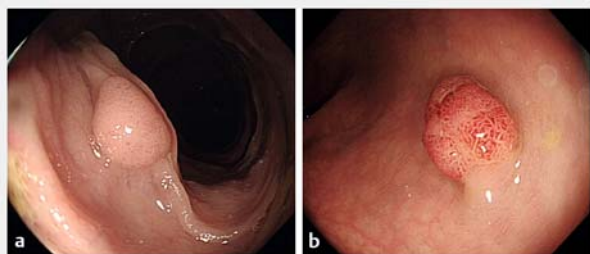
NED collects data on endoscopy processes such as indications, diagnoses, therapies, and patient comfort during each procedure, as well as anonymized data on patient age and sex. Data are recorded at the procedure level, and histology is not recorded [22].

For this retrospective cross-sectional study, within the NED-APRIQOT research project [23], we extracted data on 100 000 consecutive independent (i.e. nontraining) colonoscopies performed between 01/01/2019 and 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, urgency, therapy, polyps detected), the patient (e.g. age, sex, indication for colonoscopy), which endoscopist conducted the procedure, and the endoscopy unit/hospital Trust in which it was performed.

The dataset was restricted to complete nonemergency procedures (i.e. colonoscopy that reported extent of cecum, terminal ileum, neoterminal ileum or ileocolonic anastomosis), with withdrawal time of <120 minutes, and complete polypectomy data, conducted on patients aged 18–99 years (see **Fig. 1** in the online-only Supplementary material).

Outcome measures

We considered two recognized detection-related colonoscopy key performance indicators: mean number of polyps detected (MNP, expressed per 100 procedures, primary outcome), and polyp detection rate (PDR, secondary outcome). MNP was defined as the total number of polyps detected divided by the 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 the 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 the ADR plateaus at 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 ► **Fig. 1**.



► **Fig. 1** Examples of polyps detected during colonoscopy. a Paris Ip polyp. b 10-mm Paris Ip adenoma.

Main explanatory variables

Each colonoscopy was classified according to whether it was 1) performed on a weekend (variable: weekend Yes [performed Saturday/Sunday]/No [performed Monday–Friday]) and 2) an insourced or standard procedure (variable: insourced Yes/No [standard]). For each endoscopist, the hospital Trust where they performed most procedures during the analysis time period, was considered their main Trust. A hospital Trust may have more than one site, but an endoscopist (if they work across more than one site) would be familiar with each site, including the endoscopy equipment, team, and processes. 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 a Trust that was not the endoscopist's main Trust.

Other endoscopy 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, patient age, and procedure indication. Procedures where sex was unknown were combined with females for analysis. Patient age was categorized as 18–39, 40–49, 50–59, 60–69, and 70+ years. A hierarchical indication variable was derived: screening (indication Bowel Cancer Screening Programme or fecal occult blood test; the study was conducted before widespread use of fecal 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 computed tomography only or plus other indications but not screening, IBD or previous polyps); lower gastrointestinal symptoms (constipation, diarrhea, chronic alternating diarrhea/constipation, previous bleeding, abdominal pain, abdominal mass, or anemia only or plus other indications but not screening, IBD, previous polyps, or previous abnormal investigations); and other (polyposis, family history of colorectal cancer, colorectal cancer follow-up, stent replacement or removal, tumor assessment, weight loss or other).

Other markers of colonoscopy quality and workload [25] were considered for purposes of interpretation. Each marker was computed separately for weekend (yes/no) and insourced (yes/no) procedures. Cecal intubation rate and terminal ileum intubation rate were defined as the percentage of colonoscopies in which the colonoscope tip passed at least to the cecum and neoterminal ileum or terminal ileum, respectively. Withdrawal time was based on negative procedures only (i.e. no polyps were found and no therapeutic procedures were performed) and grouped into <6, 6–9, 10–30 minutes, and unknown. Bowel preparation quality score was categorized as excellent, good, fair, 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. Use of hyoscine butylbromide and sedation (pethidine, midazolam, or fentanyl) were also summarized.

Procedure points were used to describe the schedule of procedures (list) length, as follows. We counted and classified all of the procedures in the list that contained the index colonoscopy. Two “points” were allocated for the index colonoscopy. One additional point was allocated for each upper endoscopy or flexible sigmoidoscopy on the list; two additional points were allocated for each additional colonoscopy on the list; and three points for each endoscopic retrograde cholangiopancreatography procedure (ERCP) (e.g. had the list included only the index colonoscopy, the list length would have been “2 points”; had the list included the index colonoscopy, plus one additional colonoscopy, two flexible sigmoidoscopies and one ERCP, the list length would have been “9 points”); 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 for the NED-APRIQOT study and related analyses [23].

Statistical analysis

Descriptive statistics were calculated for the study population for case-mix, and 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 the 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 the main effects. The final case-mix adjusted weekend/weekday model was also adjusted for insourcing/standard (and vice versa). For information,

► **Table 1** Patient characteristics overall and by whether procedures were performed at weekends/weekdays, and by endoscopists working in their main institution (“standard”)/locum or agency endoscopists (“insourced”).

	Total	Weekend			Insourced		
	(N = 92 879) n (%)	Yes (N = 19 977) n (%)	No (N = 72 902) n (%)	P value ¹	Yes (N = 9909) n (%)	No (N = 82 970) n (%)	P value ¹
Age, years				0.013			<0.001
▪ 18–39	11 649 (12.5)	2367 (11.9)	9282 (12.7)		1579 (15.9)	10 070 (12.1)	
▪ 40–49	11 540 (12.4)	2522 (12.6)	9018 (12.4)		1525 (15.4)	10 015 (12.1)	
▪ 50–59	20 237 (21.8)	4377 (21.9)	15 860 (21.7)		2348 (23.7)	17 889 (21.6)	
▪ 60–69	23 461 (25.3)	5031 (25.2)	18 430 (25.3)		2166 (21.9)	21 295 (25.7)	
▪ ≥ 70	25 992 (28.0)	5680 (28.4)	20 312 (27.9)		2291 (23.1)	23 701 (28.6)	
Sex				0.001			0.072
▪ Female ²	46 861 (50.4)	10 283 (51.5)	36 578 (50.2)		5084 (51.3)	41 777 (50.4)	
▪ Male	46 018 (49.6)	9694 (48.5)	36 324 (49.8)		4825 (48.7)	41 193 (49.6)	
Indication				<0.001			<0.001
▪ Screening	8828 (9.5)	598 (3.0)	8230 (11.3)		606 (6.1)	8222 (9.9)	
▪ IBD assessment	4027 (4.3)	827 (4.1)	3155 (4.3)		377 (3.8)	3650 (4.4)	
▪ Previous polyp(s)	9760 (10.5)	2079 (10.4)	7681 (10.5)		909 (9.2)	8851 (10.7)	
▪ Previous abnormal investigation	2551 (2.8)	570 (2.9)	1981 (2.7)		276 (2.8)	2275 (2.7)	
▪ Lower GI symptoms	36 419 (39.2)	7642 (38.3)	28 777 (39.5)		4539 (45.8)	31 880 (38.4)	
▪ Other ³	31 294 (33.7)	8216 (41.1)	23 078 (32.7)		3202 (32.3)	28 092 (33.9)	

IBD, inflammatory bowel disease; GI, gastrointestinal.

¹ Chi-squared test.

² Included 1005 unknown sex.

³ Polyposis, family history of colorectal cancer (CRC), CRC follow-up, stent replacement or removal, tumor assessment, weight loss, or other.

we ran a secondary analysis fitting a 2×2 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 that we could not control for (such as family history or presence of familial syndromes), we conducted a sensitivity analysis in which we recalculated the main effects of procedure day and type, restricting consideration to procedures undertaken in patients aged 50 years 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, College Station, Texas, USA).

Results

Procedure characteristics

After exclusions, the analysis included 92 879 colonoscopies performed by 2496 endoscopists from 332 units and 112 Trusts (Fig. 1s). Overall, 53% of colonoscopies were performed in patients aged ≥60 years and half were in male patients. Investigation of lower gastrointestinal symptoms was the indication for 39.2% of colonoscopies performed, surveillance of previous

polyps accounted for 10.5%, and screening-related colonoscopies for 9.5% (► Table 1).

About one-fifth (21.5%) of colonoscopies were performed at weekends, and 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 performed on weekends, 7.6% were weekday insourced (7089), and 3.0% (2820) were weekend insourced.

At weekends, patients were less often male or attending for a screening-related procedure compared with those undergoing colonoscopy on a weekday (► Table 1). Patient age differed little between weekend and weekdays. Patients undergoing insourced colonoscopies were younger and less likely to be attending for a screening-related procedure than those receiving standard colonoscopy.

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).

► **Table 2** Mean number of polyps (MNP) per 100 colonoscopies and polyp detection rate (PDR), overall, and by whether procedures were performed at weekends or not, and were insourced¹ or not.

	Overall	Weekend		Insourced	
		Yes	No	Yes	No
MNP, mean (95 %CI) ²	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) ³	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)

¹ Insourced defined as procedures performed by locums or agency endoscopists.

² t test: weekend yes vs no, $P < 0.001$; insourced yes vs no, $P < 0.001$.

³ Chi-square tests: weekend yes vs no, $P < 0.001$; insourced yes vs no, $P < 0.001$.

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]). Similarly, 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 fully 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 years ($n = 69\,690$), 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

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

Other quality and workload indicators

► **Table 4** compares other markers of colonoscopy quality and workload by whether procedures were weekend/weekday or insourced/standard. Terminal ileum intubation rate, but not cecal intubation rate, was lower at weekends than weekdays. Two-thirds of negative weekend procedures had a withdrawal time

of < 10 minutes compared with half of weekday procedures. At weekends, a lower proportion of procedures had excellent/good bowel preparation, and hyoscine butylbromide and sedation were used less often compared with weekday procedures. Mean number of points per list was higher at weekends (10.7 [SD 3.3] vs. 8.2 [SD 3.4]).

Cecal intubation rate and terminal ileum intubation rate were both higher for insourced than standard procedures. Slightly more standard procedures had a withdrawal time of < 10 minutes. Compared with standard procedures, 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 than that for standard procedures (8.8 [SD 3.5] vs. 8.3 [SD 3.8]).

Table 2 s 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 performed by almost 2500 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, varies between weekend and weekday procedures, and between insourced and standard procedures.

The statistically significantly lower MNP and PDR for colonoscopies performed at weekends suggest a “weekend effect” on polyp detection. Independently 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 impact of strategies to increase workload, relatively few studies have examined this topic. Despite insourcing being

► Table 3 Mixed effects regression analyses of associations between weekend/weekday and insourced¹/standard procedures and mean number of polyps (MNP) per 100 colonoscopies (primary outcome) and polyp detection rate (PDR; secondary outcome).

	Univariable	Multivariable	
		Case-mix adjusted ²	Fully adjusted ³
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 ⁴			
▪ MNP, IRR (95 %CI)	0.81 (0.78–0.84) ⁵	0.87 (0.84–0.90) ⁵	0.86 (0.83–0.89) ⁵
▪ PDR, OR (95 %CI)	0.81 (0.78–0.85) ⁵	0.86 (0.82–0.90) ⁵	0.86 (0.82–0.90) ⁵
Insourced colonoscopy ⁶			
▪ MNP, IRR (95 %CI)	0.81 (0.77–0.84) ⁵	0.91 (0.87–0.95) ⁵	0.91 (0.87–0.95) ⁵
▪ PDR, OR (95 %CI)	0.76 (0.72–0.80) ⁵	0.90 (0.85–0.95) ⁵	0.90 (0.85–0.95) ⁵

IRR, incidence rate ratio; OR, odds ratio; SE, standard error.

¹ Insourced defined as procedures performed by locums or agency endoscopists.

² Adjusted for age, sex, and indication only.

³ Adjusted for age, sex, indication and, as appropriate, whether or not procedures were undertaken at weekend or were insourced. For MNP, endoscopist level variance (with standard error) was 0.27 (0.01). For PDR, endoscopist level variance was 0.39 (0.02).

⁴ Versus weekday procedure.

⁵ $P < 0.001$.

⁶ Versus standard procedure.

used in almost half of UK units [4], no previous studies appear to have examined quality of insourced colonoscopy procedures. Regarding weekend working, a single-center UK study of 17 634 patients who underwent nonscreening colonoscopies between January 2016 and November 2018 reported higher MNP and ADR during weekdays (0.49 and 28.8%, respectively) than evenings (0.38 and 24.2%, respectively) and Saturdays (0.39 and 24.4%, respectively) in univariate analyses, but these differences were not statistically significant after adjusting for working team and bowel preparation [19]. A single-center 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 indicating 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, hyoscine butylbromide use, intravenous sedation use, and points on list). The main effects for weekends 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 that 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 mindset or behaviors 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 indeed for any other future initiatives to address endoscopy capacity issues, or backlogs. Moreover, the possibility of perverse (albeit unintended) incentives to cut corners or overpopulate lists should always be considered in relation to capacity initiatives.

With regard to improving withdrawal time and hyoscine butylbromide use, reiterating quality targets is unlikely to effect

► **Table 4** Comparison of other markers of colonoscopy quality and workload, and by whether procedures were undertaken at weekends or were insourced¹.

	Weekend			Insourced		
	Yes (N = 19 977)	No (N = 72 902)	P value ²	Yes (N = 9909)	No (N = 82 970)	P value ²
CIR, % ³	93.4	93.5	0.53	94.7	93.4	<0.001
TIR, % ³	37.3	38.8	<0.001	44.6	37.7	<0.001
Withdrawal time for negative procedures ⁴ , n (%)			<0.001			0.014
▪ <6 minutes	444 (3.3)	1798 (3.9)		245 (3.7)	1997 (3.8)	
▪ 6–9 minutes	8205 (61.4)	21 357 (47.1)		3298 (49.7)	26 264 (50.4)	
▪ 10–30 minutes	3732 (28.0)	14 016 (30.9)		1827 (27.5)	15 921 (30.6)	
▪ Unknown	972 (7.3)	8203 (18.1)		1264 (19.1)	7911 (15.2)	
Bowel preparation, n (%)			<0.001			<0.001
▪ Excellent	2952 (14.8)	15 705 (21.5)		2180 (22.0)	16 477 (19.9)	
▪ Good	9852 (49.3)	35 590 (48.8)		4846 (48.9)	40 596 (48.9)	
▪ Fair	5834 (29.2)	17 398 (23.9)		2364 (23.9)	20 868 (25.1)	
▪ Inadequate	1030 (5.2)	3092 (4.2)		420 (4.2)	3702 (4.5)	
▪ Unknown	309 (1.5)	1117 (1.5)		99 (1.0)	1327 (1.6)	
Discomfort score, n (%)			<0.001			0.034
▪ None/minimal/mild	19 299 (96.6)	69 901 (95.9)		9576 (96.6)	79 624 (96.0)	
▪ Moderate/severe	512 (2.6)	2603 (3.6)		297 (3.0)	2818 (3.4)	
▪ Unknown	166 (0.8)	398 (0.5)		36 (0.4)	528 (0.6)	
Hyoscine butylbromide use, n (%)	2903 (14.5)	14 670 (20.1)	<0.001	1989 (20.1)	15 584 (18.8)	0.002
Intravenous sedation use, n (%)	13 703 (68.6)	51 905 (71.2)	<0.001	7416 (74.8)	58 192 (70.1)	<0.001
Procedure points on list, mean (SD) ⁵	10.7 (3.3)	8.2 (3.4)	<0.001	8.8 (3.5)	8.3 (3.8)	<0.001

CIR, cecal intubation rate; TIR, terminal ileum intubation rate.

¹ Insourced defined as procedures performed by locums or agency endoscopists.

² P values are by chi-square test, except for procedure points where P value is by t-test.

³ Denominator excluded four pouch procedures but otherwise was based on the full dataset without exclusions (N = 99,996).

⁴ Based on 58 727 negative procedures (weekend n = 13 353; weekdays n = 45 374; insourced n = 6634; standard n = 52 093).

⁵ Based on all 100 000 procedures.

change; information alone is generally insufficient to change behavior [32]. Instead, active interventions will be required. The Quality Improvement in Colonoscopy 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 empiric evidence on what influences colonoscopy behavior 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 experi-

ence, workload, and working environment (e.g. working in an unfamiliar team and physical environment) require investigation, and qualitative research could illuminate attitudinal, psychosocial, or experiential issues that influence behaviors and detection.

Limitations

NED does not contain histological data so we could not calculate ADR. MNP is strongly associated with ADR [24] while low PDR has been associated with increased post-colonoscopy colorectal cancer [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 the effects of initiatives to increase colonoscopy capacity on mean adenomas per patient and ADR are warranted.

Although NED is a nationwide database, not all Trusts were uploading data during the study period. However, there is no reason to believe associations between procedure day and type and detection would vary between the Trusts that were, and were not, uploading. Observed associations between procedure day and patient outcomes using routine administrative data may be an artifact of differences in quality of data recorded for weekdays and weekends [20]. NED data are uploaded in real time by direct feed from the local electronic endoscopy reporting system, completed by the endoscopist; this is likely to be a consistently reliable record of procedural findings. However, insourcing endoscopists could be less familiar with the local endoscopy reporting system, increasing risk of data entry error. Weekend/weekday procedures were identified based on procedure date entered into the endoscopy reporting system; 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).

Given the large size of the study population, care should be taken to not assume that statistical significance equates to clinical significance. Some statistically significant differences in the other quality markers 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) to be clinically meaningful and, given the association between PDR and post-colonoscopy colorectal cancer [36], likely to have implications for patient outcomes.

Our dataset comprised a combination of screening-related and other procedures. In the UK, screening-related colonoscopies 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 Bowel Cancer Screening Programme, 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 8828 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.

Acknowledgement

The authors would like to acknowledge the NED committee, Joint Advisory Group for Gastrointestinal Endoscopy (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 Dunkley; 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.

Competing interests

L. Sharp holds unrestricted project grants for endoscopy-related research from 3-D Matrix and Medtronic. L. Lu, J. Catlow, and M.D. Rutter declare that they have no conflict of interest. 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.

Funding Information

Health Foundation | AIMS ID 695428

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* 2022; 14: 103–110
- [2] Richards M. Diagnostics: recovery and renewal – report of the independent review of diagnostic services for NHS England. NHS England, 2020. Accessed: Mar 17 2022. <https://www.england.nhs.uk/publication/diagnostics-recovery-and-renewal-report-of-the-independent-review-of-diagnostic-services-for-nhs-england/>
- [3] Norgine. Scope for change: what next for European colonoscopy services. 2017. Accessed Mar 17 2023. <https://norgine.com/wp-content/uploads/2017/09/Scope-for-change-what-next-for-European-colonoscopy-services.pdf>
- [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 et al. 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: 451–460
- [7] Oats B. Getting it Right First Time (GIRFT) GIRFT Programme National Specialty Report. Accessed Mar 17 2022. https://gettingitrightfirst-time.co.uk/medical_specialties/gastroenterology/
- [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: 1795–1803

- [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: 1298–1306
- [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: E1456–E1462
- [11] Kaminski MF, 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 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: 867–874
- [16] Takagi H, Ando T, Umemoto T et al. 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: e025764
- [18] Farrow L, Hall A, Aucott L et al. 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 et al. 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 et al. Number of polyps detected is a useful indicator of quality of clinical colonoscopy. *Endosc Int Open* 2018; 6: E878–884
- [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 et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. *Gut* 2012; 61: 1050–1057
- [27] Brown H, Prescott R. *Statistics in Practice. Applied mixed models in medicine.* West Sussex UK: John Wiley & Sons; 2006
- [28] Dong Z, Wang J, Chen Y et al. Negative effects of endoscopists' fatigue on colonoscopy quality on 34,022 screening colonoscopies. *J Gastrointest Liver Dis* 2021; 30: 358–365
- [29] Teng TY, Khor SN, Kailasam M et al. Morning colonoscopies are associated with improved adenoma detection rates. *Surg Endosc* 2016; 30: 1796–1803
- [30] Marcondes FO, Gourevitch RA, Schoen RE et al. Adenoma detection rate falls at the end of the day in a large multi-site sample. *Dig Dis Sci* 2018; 63: 856–859
- [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: 203–211
- [32] Kelly MP, Barker M. Why is changing health-related behaviour so difficult? *Public Health* 2016; 136: 109–116
- [33] Neilson LJ, East JE, Rajasekhar PT et al. Sustained colonoscopy quality improvement using a simple intervention bundle. *Endoscopy* 2020; 52: 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: 1030–1040.e
- [35] Catlow J, Bhardwaj-Gosling R, Sharp L et al. 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: 704–715
- [36] Schwarz S, Hornschuch M, Pox C et al. Polyp detection rate and cumulative incidence of post-colonoscopy colorectal cancer in Germany. *Int J Cancer* 2022; 152: 1547–1555