

# Effect of quality metric monitoring and colonoscopy performance

## Authors

Anthony Razzak<sup>1</sup>, Dineen Smith<sup>2</sup>, Maliha Zahid<sup>3</sup>, Georgios Papachristou<sup>1</sup>, Asif Khalid<sup>1</sup>

## Institutions

<sup>1</sup> University of Pittsburgh Medical Center, Pittsburgh, PA, USA

<sup>2</sup> VA Pittsburgh Health Care System, Pittsburgh, PA, USA

<sup>3</sup> University of Pittsburgh, Pittsburgh, PA, USA

## submitted

11. November 2015

## accepted after revision

13. June 2016

## Bibliography

DOI <http://dx.doi.org/10.1055/s-0042-110787>  
Published online: 5.8.2016  
Endoscopy International Open 2016; 04: E1023–E1027  
© Georg Thieme Verlag KG  
Stuttgart · New York  
E-ISSN 2196-9736

## Corresponding author

Asif Khalid, MD

M2  
C-wing  
PUH  
200 Lothrop Street  
Pittsburgh  
PA 15213  
USA  
Fax: +1-412-383-8992  
khalida@upmc.edu

**Background and aims:** Adenoma detection rate (ADR) and cecal withdrawal time (CWT) have been identified as measures of colonoscopy quality. This study evaluates the impact of monitoring these measures on provider performance.

**Methods:** Six blinded gastroenterologists practicing at a Veterans Affairs Medical Center were prospectively monitored over 9 months. Data for screening, adenoma surveillance, and fecal occult blood test positive (FOBT+) indicated colonoscopies were obtained, including exam preparation quality, cecal intubation rate, CWT, ADR, adenomas per colonoscopy (APC), and adverse events. Metrics were continuously monitored after a period of informed CWT monitoring and informed CWT+ADR monitoring. The primary outcome was impact on ADR and APC.

**Results:** A total of 1671 colonoscopies were performed during the study period with 540 before informed monitoring, 528 during informed CWT monitoring, and 603 during informed CWT+ADR

monitoring. No statistically significant impact on ADR was noted across each study phase. Multivariate regression revealed a trend towards fewer adenomas removed during the CWT monitoring phase (OR=0.79; 95%CI 0.62–1.02,  $P=0.065$ ) and a trend towards more adenomas removed during the CWT+ADR monitoring phase when compared to baseline (OR=1.26; 95%CI 0.99–1.61,  $P=0.062$ ). Indication for examination and provider were significant predictors for higher APC. Provider-specific data demonstrated a direct relationship between high ADR performers and increased CWT.

**Conclusions:** Monitoring quality metrics did not significantly alter colonoscopy performance across a small heterogeneous group of providers. Non-significant trends towards higher APC were noted with CWT+ADR monitoring. Providers with a longer CWT had a higher ADR. Further studies are needed to determine the impact of monitoring on colonoscopy performance.

## Introduction

Colonoscopy is associated with a decrease in colorectal cancer incidence and colorectal cancer related mortality [1–3]. Variability in colonoscopy performance has been widely reported and is a focus of interest when identifying quality metrics and optimizing colorectal cancer prevention [4–6]. Both adenoma detection rate (ADR) and cecal withdrawal time (CWT) are recognized as measures of colonoscopy quality [6]. An increased CWT has been associated with increased adenoma detection [4–7]. ADR, defined as the proportion of screened subjects with at least one histologically confirmed adenomatous lesion detected, is inversely associated with interval colorectal cancer and colorectal cancer related mortality [8–10]. Current guidelines recommend an ADR  $\geq 25\%$  and  $\geq 15\%$  in men and women, respectively [6]. Owing to its increasingly recognized impor-

tance, the Centers for Medicare and Medicaid Services adopted ADR as a Physician Quality Reporting System measure in 2014, and it has been estimated that colonoscopy reimbursement may soon be directly associated with quality metric targets.

## Background

Multiple efforts have been made to identify interventions that beneficially impact adenoma detection. Improved adenoma detection has been demonstrated with performance-based interventions including optimized CWTs [4, 11], cecal retroflexion with retroflexed proximal colonic evaluation [12], and implementation of a split dose colon cleanse [13]. However, the impact of monitoring, feedback, and educational programs has demonstrated mixed results likely related to the hetero-

## License terms



geneity of intervention, participants, and practice environment [5, 14–16]. Most recently, the endoscopic quality improvement project (EQUIP) demonstrated sustained improvement in ADR via implementation of two 1-hour educational programs, access to educational material, and monthly ADR feedback [17, 18]. The impact was similar when adenomas per patient were analyzed. If the goal of colonoscopy is to decrease colorectal cancer incidence, then the aim for the endoscopist is surely to remove all adenomas. A higher CWT is a surrogate for a more detailed exam with increased ADR. The optimum CWT to maximize ADR however, remains elusive. Monitoring of CWT will provide a goal for endoscopists in minutes. What is unclear is whether “making” an endoscopist increase his/her average CWT will lead to improvement in his/her ADR or APC (adenoma per colonoscopy). Data suggest that endoscopists with a CWT greater than 6 minutes have a higher ADR. We do not know if taking endoscopists with a CWT less than 6 minutes and intervening to increase their CWT to over 6 minutes will increase their ADR proportionately. Likewise, monitoring ADR provides a target for the endoscopist. The target is reached when one adenoma is removed. We know a higher ADR is associated with decreased interval colorectal cancer and colorectal cancer related mortality. Questions that remain unknown are (1) will monitoring of ADR increase endoscopist ADR from the baseline before monitoring, and (2) will ADR monitoring also increase APC. This study aims to evaluate the impact of quality metric monitoring on colonoscopy performance.

## Methods

This is a prospective study performed at a large tertiary care VA hospital in Pittsburgh. Patients are scheduled for outpatient colonoscopy at our institution through the following mechanisms:

1. Direct to gastrointestinal lab through an open access system.
2. Scheduled after being evaluated in the outpatient gastrointestinal clinic.
3. Scheduled for follow-up/surveillance at the time of previous colonoscopy.

Colonoscopies are randomly assigned to physician rooms. All patients were offered a standard GoLYTELY® (Braintree Laboratories Inc., MA, United States) preparation over the study period, unless there was a history of preparation inadequacy or intolerance to GoLYTELY. Patients were asked to stay on a clear liquid diet and drink 8 ounces of GoLYTELY every 10 minutes starting at 1700 the day before colonoscopy until consumed completely.

Colonoscopy exams included those performed for colorectal cancer screening, surveillance (history of prior adenoma) or positive fecal occult blood test (FOBT) only. Exams performed for other indications or with a trainee were excluded. Data collection included patient demographics, exam details (indication, time of exam, preparation quality, cecal intubation rate, findings, polyp details [number seen, number removed, size, location, and pathology], and adverse events), and quality measures. The quality measures included CWT (in negative screening exams), ADR (proportion of screening exams with at least one adenoma), and APC (adenomas per colonoscopy). This study was performed in three phases.

Phase 1: This comprised obtaining the baseline colonoscopy data of six gastroenterologists over a time period from the year previous to the study conception date. The time frame required to reach a minimum of 500 cases (see statistical analysis below)

was 3 months. All gastroenterologists excluding the principal investigator were unaware of this data collection.

Phase 2: An email memorandum was sent to all of the gastroenterologists by the gastrointestinal section chief that documentation of the CWT was now required as a quality measure. The process was defined as follows: Once the cecum was identified by photo-documenting the appendiceal orifice and ileocecal valve, the assisting registered nurse (RN) will be asked to turn on a timer. The timer will be turned off upon withdrawal of the scope from the rectum, and the time noted in the templated report. The timer will not be turned off for periods of water lavage, polyp removal or mucosal sampling. The CWT will only be calculated for exams in which no samples were collected. No expected or preferred target CWT was provided to the gastroenterologists. Data were collected for 3 months as described in Phase 1, now including CWT. All gastroenterologists excluding the principal investigator were unaware of this data collection.

Phase 3: An email memorandum was sent to all of the gastroenterologists by the gastrointestinal section chief that individual ADR would be monitored as a quality measure. The gastroenterologists were informed that the individual ADR information for each gastroenterologist would be kept private but would be linked to a performance incentive awarded to those who achieved an ADR >25%. The gastroenterologists were not informed of the timing of ADR calculation. Data were collected for 3 months as described in Phase 2. All gastroenterologists excluding the principal investigator were unaware of this data collection.

Data entry and statistical analyses were performed by DS and MZ, respectively. None of the gastroenterologists contributed to the data entry or statistical analysis.

The VAPHS IRB approved this study. At the time of study renewal 1 year later, the study was put on hold due to the concern that the gastroenterologists had not been identified as research subjects. Only the patients had been identified as research subjects. At this time, the study was in data compilation and analysis phase. After additional review, the study was given approval provided the gastroenterologists gave informed consent as research subjects, which was obtained.

## Statistical analysis

The duration of each study phase was determined by a sample size calculation based on a minimum desired ADR of 25% for the group. To detect a 10% change in ADR (to 27.5%) with an alpha value of 0.05 and power of 80%, at least 252 patients would be needed in each phase of the study. Given (1) the ADR of the group was unknown and may have been dramatically different from 25%, and (2) the relative contribution of cases per provider may vary based on vacation, etc., we arbitrarily doubled the target sample size for each phase to 500. In the first phase, we were able to reach the desired 500 cases (542) in a 3-month period. As such, we decided to have each phase be at least 3 months in length and at least yield 500 cases.

All continuous variables are presented as mean  $\pm$  standard deviation. Continuous variables across two groups were compared using the Student's *t* test with one-way analysis of variance applied across multiple groups with post-hoc Bonferroni correction applied for multiple testing. Categorical variables were compared using the Chi-square test, with Fisher exact correction applied if any cell variable was 5 or less. Univariate logistic regression was used with adenoma detection as the outcome variable of interest. All independent variables with a *P* value <0.10 were entered in a

	All patients	Phase 1	Phase 2	Phase 3
No. of patients	1671	540	528	603
Age (mean/median) [SD]	61.2/62 [7.3]	61.3 [7.2]	61 [7]	61.3 [7.7]
Sex (% males)	95.30	95.56	95.45	95.02
Indication, % ( $P < 0.001$ )				
Screening	51.10	43.50	56.40	53.00
Surveillance	8.90	11.50	7.40	8
FOBT+	40	45	36.20	39
Complete exam, %	99.60	99.26	100	99.50
Exam preparation, % ( $P = 0.005$ )				
Adequate	53.80	55.56	56.82	49.59
Fair	44.20	43.70	41.29	47.10
Poor	2	0.74	1.89	3.32
Exam time (AM %)	74.10	71.70	74.20	76.10

FOBT+, positive fecal occult blood test; SD, standard deviation.

step-up multivariate logistic regression model. A two-tailed  $P$  value of  $< 0.05$  was considered significant. All analyses were performed using Stata 12.0.

## Results

A total of 1671 colonoscopies were performed during the study period with 540 in phase 1, 528 in phase 2, and 603 during phase 3. The exams were performed with air insufflation, and water lavage was used at the discretion of the gastroenterologist. **Table 1** summarizes the data collected by phase of study. The three phases of the study were significantly different in a number of ways:

- Exam preparation: The exam preparation was reported as inadequate in a significantly higher number of patients in phase 3 [ $P = 0.005$ ] (**Table 1**) for the doctors as a group and this difference reached significance for doctor 6 individually [ $P = 0.019$ ] (**Table 2**). There was wide variation in the reported exam preparation adequacy between the doctors [ $P < 0.001$ ] (**Table 3**).
- Exam indication: There were significantly more patients undergoing colonoscopy for FOBT+ and surveillance indications [ $P < 0.001$ ] in group 1. APC was highest for FOBT indicated procedures (61.1), followed by surveillance (54.7), and

screening (45.6) colonoscopies. Thus, group 1 was enriched with patients more likely to have adenomas (**Table 1** and **Table 2**).

- The relative contribution of number of exams from each doctor for the study varied significantly [ $P = 0.003$ ] (**Table 3**). Although the numbers are small, variation was also present in the relative number of exams per phase of study from each doctor.

In analyses uncontrolled for these variables, during phase 3 of the study, there were significantly higher mean number of polyps  $< 5$  mm seen per exam, mean number of polyps  $< 10$  mm removed per exam, exams with an adenoma  $< 5$  mm in size, and exams with a right-sided adenoma. This is in spite of group 1 being relatively enriched with patients with FOBT+ and surveillance indications, and group 3 having the least number of patients with an adequate exam. The ADR was not significantly different, however. A trend was seen towards a higher rate of post-polypectomy bleeding in group 3.

In a multivariate regression model, a trend was seen towards fewer adenomas removed per exam when CWT was being monitored [OR = 0.79; 95%CI 0.62 – 1.02,  $P = 0.065$ ] and more adenomas removed per exam when ADR was being monitored [OR = 1.26; 95%CI 0.99 – 1.61,  $P = 0.062$ ] (**Table 4**). Significant predictors of a higher number of adenomas being removed per exam included the indication for the exam and which doctor per-

**Table 2** Overall results of the study and by phase of the study.

	All patients	Phase (P) 1	Phase (P) 2	Phase (P) 3
Mean [SD] CWT, min ( $P = ns$ )	9.6 [2.6]		9.7 [2.9]	9.4 [2.2]
Exams with a polyp seen, n (%)	1089 (65.2)	348 (64)	328 (62.2)	413 (68.3)
Mean [SD] no. polyps seen per exam ( $P < 0.001$ ) ( $P = 0.02$ P1&3, $P = 0.08$ P2&3)	1.9 [2.6]	1.7 [2.2]	1.8 [2.4]	2.1 [3.1]
Mean [SD] no. of polyps removed per exam ( $P < 0.001$ ) ( $P = 0.04$ P1&3, $P = 0.058$ P2&3)	1.8 [2.4]	1.7 [2.0]	1.7 [2.3]	2 [2.7]
Mean [SD] no. of polyps $< 5$ mm removed per exam ( $P < 0.001$ ) ( $P = 0.049$ P1&3, $P = 0.03$ P2&3)	1.2 [1.8]	1.14 [1.5]	1.12 [1.64]	1.39 [2]
Mean [SD] no. of polyps 5 – 9 mm removed per exam ( $P < 0.001$ )	0.47 [1]	0.47 [1.1]	0.47 [1]	0.51 [1]
Exams with a polyp $> 9$ mm removed, n (%)	137 (8)	50 (9)	34 (6)	53 (9)
Mean [SD] no. of polyps $> 9$ mm removed per exam ( $P < 0.001$ )	0.1 [0.5]	0.14 [0.7]	0.09 [0.5]	0.1 [0.4]
Exams with an adenoma ( $P = 0.006$ P1&2, $P < 0.001$ P1&3), n (%)	848 (51)	281 (55)	230 (45)	337 (54.5)
Exams with a right colonic adenoma ( $P = 0.09$ ) (OR = 0.7 P2&3, $P = 0.01$ ), n (%)	394 (24)	130 (24)	106 (20)	158 (26)
Mean [SD] no. of adenomas ( $P = 0.005$ P1&3, $P < 0.001$ P2&3)	1.2 [1.9]	1.2 [1.7]	1 [1.8]	1.4 [2.1]
Mean [SD] no. of adenomas $< 5$ mm removed per exam ( $P = 0.001$ P2&3)	0.75 [1.4]	0.74 [1.2]	0.6 [1.2]	0.89 [1.6]
Mean [SD] no. of adenomas 5 – 9 mm removed per exam	0.34 [0.84]	0.29 [0.63]	0.34 [1]	0.39 [0.9]
Mean [SD] no. of adenomas $> 9$ mm or TVA removed per exam	0.1 [0.5]	0.12 [0.5]	0.06 [0.3]	0.11 [0.5]
ADR ( $P = ns$ ), %	45.60	46.38	42.62	47.81

ADR, adenoma detection rate; CWT, cecal withdrawal time; OR, odds ratio; SD, standard deviation; ns, not significant.

**Table 3** Data for each doctor.

Doctor	Number of patients ( $P=0.003$ )	Adequate exam prep, % ( $P<0.001$ )	Mean CWT, min	% Exams with a polyp ( $P<0.001$ )	% Exams with an adenoma ( $P<0.001$ )	ADR, % ( $P<0.001$ )	OR for adenoma (controlled) ( $P<0.001$ )
1	295	75	8.8	57.6	38	32.7	1
2	242	55	8.4	47.5	36	31.4	0.92
3	248	34	11.9	66.9	58.5	59.4	2.3
4	302	51	8.9	66.2	50.7	44.2	1.7
5	359	63	12.3	80.8	66	59.3	3.2
6	225	34	9.9	65.8	49.8	48.3	1.6

ADR, adenoma detection rate; CWT, cecal withdrawal time; OR, odds ratio.

formed the exam (Table 4). Overall, no statistically significant impact on ADR was noted across each study phase. Provider-specific data demonstrated heterogeneity across all providers with a direct relationship between high ADR performers and increased CWT (Table 3), regardless of intervention.

## Discussion

In this study, no significant change in colonoscopy performance was noted during quality metric monitoring. A non-significant trend towards fewer adenomas detected during the period of CWT monitoring and a non-significant trend towards more adenomas detected while monitoring CWT+ADR was identified. More bleeding complications and diminutive polyps were identified in the CWT+ADR monitored phase; however, this did not appear significant when controlled for provider and indication for examination. Inter-provider evaluation revealed significant heterogeneity across the providers with ADRs ranging from 31.4 to 59.4. A near direct correlation between ADR and CWT was noted, with high-ADR performers having the highest CWT. During the CWT and ADR monitoring phase, three of the six providers endorsed a noticeable reduction in the number of colon preparations identified as adequate.

At face value, the data herein suggest that monitoring quality metrics has little impact on the performance of colonoscopy. Certain findings however, could raise suspicion that mandated colonoscopic diligence could result in unwanted behavioral changes. For instance, does a more thorough examination result in removal of more diminutive and non-adenomatous polyps that have questionable clinical significance and not impact colorectal cancer risk in future? And does this action predispose to an increase in post-polypectomy related adverse events? While not signifi-

cant, our study demonstrated an increase in both factors when CWT+ADR monitoring was employed. In addition, half of the providers reported worse colon preparations during the CWT+ADR monitoring phases despite no change in the patient population or colon preparation utilized. Will the threshold to identify an "inadequate" preparation change if providers are more or less successful at detecting adenomas? Will this result in more frequent, possibly unnecessary, follow-up procedures? Recent data have demonstrated that colon preparation is a factor in inconsistent colonoscopy follow-up recommendations relative to current guidelines [19].

A recent retrospective study evaluating the impact of reporting quality indicators, including CWT and ADR, on colonoscopy performance revealed a higher adenoma detection rate, 53.9% (95% CI-49.7–58.1%) vs. 44.7% (95%CI-39.1–50.4%),  $P=0.13$  after receiving feedback with the incremental increase in ADR mostly due to an increase in proximal adenoma detection. Despite the intervention, there was no statistically significant impact on adenomas per colonoscopy (APC), serrated polyp detection, advanced neoplasm identification or mean size of adenoma [16]. Current quality metrics appear to be appropriate surrogates for high quality colonoscopists, but interventions to impact low quality performers are likely complex and difficult to implement. It is probable that high quality endoscopists share specific traits that translate to high ADR and high CWT, but are not defined by such measures. For instance, the obsessive, detail oriented provider who spends time looking behind each haustra, lavage and suction of stool, is likely motivated by factors beyond quality metrics and has specific character traits that prompt such behavior. Mandating specific quality metric targets may not result in improved performance without concomitant programs to adjust behavior in poor performing providers. The success of the EQUIP trial may begin to highlight a non-performance-based interven-

		Odds Ratio*	P value	95% confidence interval
Indication	Screening	1		
	FOBT+	1.84	0.001	1.27–2.67
	Surveillance	1.39	0.002	1.13–1.72
Study phase	Phase 1	1		
	Phase 2	0.79	0.065	0.61–1.01
	Phase 3	1.26	0.062	0.99–1.6
Doctor	1	1		
	2	0.90	0.564	0.63–1.29
	3	2.21	<0.001	1.56–3.14
	4	1.67	0.002	1.2–2.32
	5	3.2	<0.001	2.32–4.43
	6	1.58	0.011	1.11–2.26

FOBT+, positive fecal occult blood test.

\* Odds ratio of an adenoma being removed per colonoscopy.

**Table 4** Variables independently (significant/not significant but with a trend) associated with APC (number of adenomas removed per exam) in multivariate regression model.

tion that can provide such impact [17, 18]. But the question will persist until ample evidence is available, can a low-ADR performer become a high-ADR provider.

In summary, monitoring quality metrics did not appear to significantly alter colonoscopy performance across a small heterogeneous group of providers. Non-significant trends towards higher APC were noted with CWT+ADR monitoring. Provider-specific data demonstrated direct associations with high ADR performance and increased CWT. More studies are needed to determine the impact of monitoring on colonoscopy performance.

**Competing interests:** None

## References

- 1 Zauber AG, Winawer SJ, O'Brien MJ et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *NEJM* 2012; 366: 687–696
- 2 Nishihara R, Wu K, Lochhead P et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *NEJM* 2013; 369: 1095–1105
- 3 Baxter NN, Warren JL, Barrett MJ et al. Association between colonoscopy and colorectal cancer mortality in a US cohort according to site of cancer and colonoscopist specialty. *J Clin Oncol* 2012; 30: 2664–2669
- 4 Barclay RL, Vicari JJ, Doughty AS et al. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *NEJM* 2006; 355: 2533–2541
- 5 Corley DA, Jensen CD, Marks AR. Can we improve adenoma detection rates? A systematic review of intervention studies *Gastrointest Endosc* 2011; 74: 656–665
- 6 Rex DK, Petrini JL, Baron TH et al. Quality indicators for colonoscopy. *Am J Gastroenterol* 2006; 101: 873–885
- 7 Lee TJW, Blanks RG, Rees CJ et al. Longer mean colonoscopy withdrawal time is associated with increased adenoma detection: evidence from the Bowel Cancer Screening Programme in England. *Endoscopy* 2013; 45: 20–26
- 8 Kaminski MF, Regula J, Kraszewska E et al. Quality indicators for colonoscopy and the risk of interval cancer. *NEJM* 2010; 362: 1795–1803
- 9 Zhao WK, Lee JK, Doubeni CA et al. Adenoma detection rate and risk of colorectal cancer and death. *NEJM* 2014; 370: 1298–1306
- 10 Baxter NN, Sutradhar R, Forbes SS et al. Analysis of administrative data finds endoscopist quality measures associated with postcolonoscopy colorectal cancer. *Gastroenterology* 2011; 140: 65–72
- 11 Barclay RL, Vicari JJ, Greenlaw RL. Effect of a time-dependent colonoscopic withdrawal protocol on adenoma detection during screening colonoscopy. *Clin Gastroenterol Hepatol* 2008; 6: 1091–1098
- 12 Hewett DG, Rex DK. Miss rate of right-sided colon examination during colonoscopy defined by retroflexion: an observational study. *Gastrointest Endosc* 2011; 74: 246–252
- 13 Gurudu SR, Ramirez FC, Harrison ME et al. Increased adenoma detection rate with system-wide implementation of a split-dose preparation for colonoscopy. *Gastrointest Endosc* 2012; 76: 603–608
- 14 Shaikat A, Oancea C, Bond JH et al. Variation in detection of adenomas and polyps by colonoscopy and change over time with a performance improvement program. *Clin Gastroenterol Hepatol* 2009; 7: 1335–1340
- 15 Lin OS, Kozarek RA, Arai A et al. The effect of periodic monitoring and feedback on screening colonoscopy withdrawal times, polyp detection rates, and patient satisfaction scores. *Gastrointest Endosc* 2010; 71: 1253–1259
- 16 Kahi CJ, Ballard D, Shah AS et al. Impact of a quarterly report card on colonoscopy quality measures. *Gastrointest Endosc* 2013; 77: 925–931
- 17 Ussui V, Coe S, Rizk C et al. Stability of increased adenoma detection at colonoscopy. Follow-up of an endoscopic quality improvement program – EQUIP-II. *Am J Gastroenterol* 2015; 110: 489–496
- 18 Coe SG, Crook JE, Diehl NN et al. An endoscopic quality improvement program improves detection of colorectal adenomas. *Am J Gastroenterol* 2013; 108: 219–226
- 19 Menees SB, Elliott E, Govani S et al. The impact of bowel cleansing on follow-up recommendations in average-risk patients with a normal colonoscopy. *Am J Gastroenterol* 2014; 109: 148–154