

Long-Term Assessment of the Cecal Intubation Rates in High-Performing Colonoscopists: Time for Review

Krishna C. Vemulapalli, MBBS, MPH¹, Sarah W. Wilder, BA¹, Charles J. Kahi, MD, MS¹ and Douglas K. Rex, MD¹

- OBJECTIVES:** The cecal intubation rate (CIR) is one of the 3 priority indicators for quality in colonoscopy. Whether continuous measurement of CIR is useful in high performers is uncertain.
- METHODS:** At an academic center, we identified 16 physicians who performed at least 50 procedures over 6 consecutive years. We analyzed all colonoscopy procedures excluding those with poor/inadequate preparation or severe colitis for CIR trend over the years. We calculated the numbers needed to establish CIR over minimum threshold levels with 95% confidence.
- RESULTS:** The overall CIR was 99.4%. None of the 16 physicians had a CIR <96.6% in any year. Sensitivity analyses including patients without intent to reach the cecum and inadequate bowel preparation had little impact on the results. Overall cecal photo documentation rate was 98.4%. No significant correlation was observed between procedure volume at our center and CIR ($\sigma = -0.196$, $P = 0.483$). Physicians with CIR $\geq 99\%$ need to have only 24 examinations reviewed to establish CIR is $>95\%$.
- DISCUSSION:** Continuous measurement of CIR, at least in high performers, appears to be of limited value. Very high performers need to evaluate small number of cases to demonstrate that CIR is above the recommended thresholds.

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/CTG/A231>, <http://links.lww.com/CTG/A232>, <http://links.lww.com/CTG/A233>, <http://links.lww.com/CTG/A234>, <http://links.lww.com/CTG/A235>, <http://links.lww.com/CTG/A236>

Clinical and Translational Gastroenterology 2020;11:e00153. <https://doi.org/10.14309/ctg.000000000000153>

INTRODUCTION

In 2015, because past recommendations to measure quality in a large number of areas were beyond the feasibility of many endoscopy units, the American College of Gastroenterology/American Society for Gastrointestinal Endoscopy (ACG/ASGE) Task Force on quality in endoscopy named 3 “priority” quality indicators for colonoscopy, including the adenoma detection rate (ADR), cecal intubation rate (CIR), and the use of appropriate screening and surveillance intervals (1).

Recent trends in quality measurement in colonoscopy are to focus on priority quality indicators and to measure them continuously over years. However, our observation has been that individual physician CIRs (2) and ADRs (3) tend to remain stable or increase over time, but rarely decline. In centers with electronic report generating systems, calculation of CIR is simple. In centers without electronic systems, calculation of CIR requires a manual review of records which is time consuming and expensive. In all centers, an audit of cecal photography requires manual record reviews. Given these costs, and because CIRs are not observed to decline, continuous measurement of these quality indicators for

individuals well above these thresholds could have a limited value. Reducing effort and resources applied to the measurement of quality indicators for individuals with consistently high performance could free resources to measure other quality indicators. For example, recent evidence suggests that polyp resection skills are not correlated with detection skills (4). Thus, individuals with high CIRs and ADRs could be poor performers with regard to resection.

In this study, we investigated the stability of CIRs over time. First, we examined whether CIRs declined over time and whether they declined below the minimum thresholds. Second, we calculated the minimum number of colonoscopies needed to assess whether colonoscopists with various levels of CIRs were remaining above the minimum recommended thresholds. Although current recommendations on quality stipulate that CIR should be $\geq 90\%$ overall for all colonoscopies (5) and $\geq 95\%$ for screening colonoscopy (6), our experience is that gastroenterologists achieve CIRs for all colonoscopies that are consistently $\geq 95\%$ and we used 95% as the minimum acceptable threshold for all colonoscopies. Finally, we tested to see whether there is an association between CIR and number of procedures performed.

¹Division of Gastroenterology and Hepatology, Indiana University School of Medicine, Indianapolis, Indiana, USA. **Correspondence:** Douglas K. Rex, MD. E-mail: drex@iu.edu.

Received September 30, 2019; accepted February 13, 2020; published online March 17, 2020

© 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

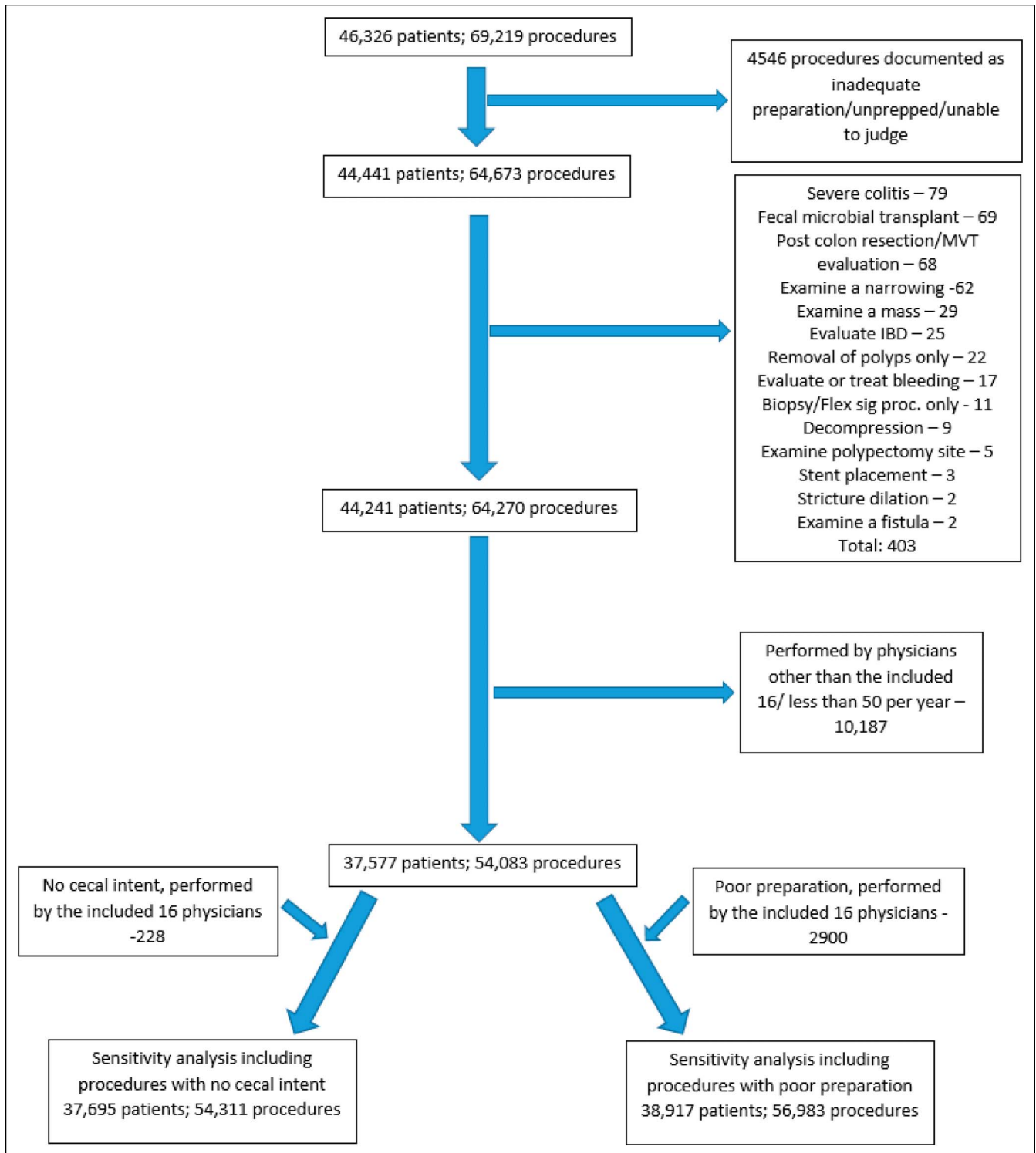


Figure 1. Patient flow through the eligibility criteria.

Recommendations regarding whether certain patients should or can be excluded from the CIR calculations vary between the guideline groups. Thus, the ACG/ASGE quality task force recommends cases need not be counted if they were aborted for inadequate preparation or if there was no intent to reach the cecum (e.g., treatment of a stricture or large polyp in the distal colon). The English National Health Service (7) and European Society for Gastrointestinal

Endoscopy (5) recommend an unadjusted CIR of $\geq 90\%$ in screening colonoscopies, and the European commission recommends a 90% minimum CIR after exclusion of patients with obstructing cancer requiring surgery (8). To increase the generalizability of our results, we performed analyses with and without inclusion of patients with inadequate preparation and also where the stated purpose was to perform a treatment without intent to intubate the cecum.

Table 1. Colonoscopist procedure volume by year

Colonoscopist	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total
1	0	0	0	208	196	235	215	353	317	232	115	0	0	0	0	0	0	1,871
2	0	0	0	0	0	0	0	58	86	68	108	123	86	100	112	105	102	948
3	0	0	0	0	0	59	96	148	105	51	93	88	50	0	0	0	0	690
4	0	0	0	0	0	0	0	0	0	109	222	293	395	423	544	563	557	3,106
5	70	235	171	216	242	238	250	363	422	418	408	424	340	221	0	0	0	4,018
6	0	0	0	0	0	0	0	0	0	0	217	310	302	311	423	536	599	2,698
7	0	0	0	0	0	0	0	123	91	68	75	118	71	64	75	96	102	883
8	0	0	0	0	167	255	311	237	270	269	317	361	275	307	353	450	399	3,971
9	0	0	0	0	0	0	0	67	67	78	71	100	101	98	100	0	0	682
10	0	0	0	0	0	0	0	144	171	182	156	178	152	182	297	384	392	2,238
11	0	0	0	0	0	0	0	0	0	0	71	274	251	269	298	224	176	1,563
12	321	1,449	1,305	1,235	1,061	1,000	1,008	1,278	1,319	1,328	1,512	1,576	1,539	1,680	1,966	1,993	2,010	23,580
13	0	0	0	0	0	0	0	0	0	82	342	339	310	344	407	350	422	2,596
14	0	0	0	0	0	0	0	0	0	0	0	136	296	337	396	491	554	2,210
15	0	0	0	0	0	156	173	154	209	107	115	142	135	150	143	129	66	1,679
16	0	0	0	0	0	0	0	0	0	68	168	180	182	239	211	188	114	1,350
Total	391	1,684	1,476	1,659	1,666	1,943	2,053	2,925	3,057	3,060	3,990	4,642	4,485	4,725	5,325	5,509	5,493	54,083

METHODS

We used data from the endoscopic quality programs at Indiana University Hospital and its associated outpatient endoscopy units. From the cohort of gastroenterology attendings at Indiana University, we identified 16 physicians who performed at least 50 colonoscopies per year for 6 consecutive years. Data were extracted directly from ProVation (ProVation Medical, Minneapolis, MN) for the years 2002–2018 which included the demographic data for patients with the date of examination, performing physician, location reached according to the colonoscopy report, and bowel preparation quality. We chose 2002 because that was the year ProVation was fully introduced and 50 minimum colonoscopies as previous studies have used that number before (9,10). In our center, we define cecal intubation as fully intubating the cecal caput, with identification of the appendiceal orifice, and or the terminal ileum, and allowing full examination of the medial wall of the cecum. Endoscopists identify each cecal landmark and note each in the colonoscopy report. Patients with right colon resections were included in the study and counted as full colonoscopy if the anastomosis was reached and noted in the report. Patients in whom the examination was aborted because of bowel preparation quality rated as poor or inadequate and those who had severe colitis were excluded from the primary analysis. Because it was not always clear in the cases of inadequate preparation whether the procedure was aborted in some regard because of the preparation, we performed a sensitivity analysis adding all patients with inadequate preparation. Patients documented as no cecal intent were excluded from the study (1), although we performed a sensitivity analysis including these colonoscopies to repeat the primary analysis of variation in cecal intubation over the years as a whole and by each physician. Throughout the study, propofol was the primary sedative agent.

Documentation by photography was evaluated as a separate issue from notation of cecal intubation and notation of cecal landmarks. We performed a manual audit of 24 randomly selected consecutive

procedures of each physician to determine the percent of cecal intubations for which cecal photography was obtained. The cecum was considered photographed if there was a convincing photograph of either the appendiceal orifice, terminal ileum, and/or the cecum from just distal to the ileocecal valve, although most cases included photography of each area. Permission to perform this study was granted by the Institutional Review Board at Indiana University.

Statistical analysis

We present the data for cecal intubation as a line graph with each line representing a single physician. As a single subject repeated measures design, any quantitative analysis for trend is not appropriate and visual interpretation is presented instead. The necessary number of procedures needed to review for a consistently high level of performance was calculated based on the level of performance compared with the recommended minimum threshold. For example, an endoscopist with CIR consistently above 99% would require a smaller number of examinations reviewed for a 95% confidence interval that performance was continuing above the 95% CIR level, compared with an endoscopist with a 98%, vs 97%, etc. The number of cases with audit at photography (n = 24) was also based on these calculations. We used the following formula: $n = \lceil \frac{z^2 p(1-p)}{e^2} \rceil$ where z = z-score corresponding to the confidence level, p is the sample proportion, and e is the margin of error. Spearman correlation was used to test the association between the number of procedures and CIR. All analyses were performed using SPSS Statistics, version 25 (IBM Corp, Armonk, NY).

RESULTS

During the study period (September 2002 to December 2018) 46,326 patients underwent 69,219 procedures. Figure 1 shows the flow of patients through the eligibility criteria and number of patients excluded with reasons. A total of 37,577 patients (56% women, average age: 55.6 years) undergoing 54,083 procedures

Table 2. Reasons for incomplete examination in the 313 colonoscopies

Reason for incomplete examination	No. of procedures, n (%)
Redundant/dilated colon or looping	91 (29.1)
Luminal narrowing	77 (24.6)
Restricted mobility	42 (13.4)
Obstructing mass	41 (13.1)
Sigmoid angulation or severe diverticulosis	25 (8)
Sedation issues	14 (4.5)
Hernia/fistula	9 (2.9)
Not able to ascertain a reason	9 (2.9)
Perforation	5 (1.6)

remained available for the final analysis. Sixteen attending gastroenterologists with varying number of years in practice completed these procedures. The overall CIR was 99.4%. The year where colonoscopies first appear should not be considered the year of onset of practice for individual physicians because physicians were often attending at our county hospital or Veterans Administration Hospital within our medical center before they began performing procedures at the study endoscopy center.

Table 1 shows the number of procedures performed per year by a physician. Table 2 provides the reasons for an incomplete examination in 313 procedures where the cecum was not intubated. Figure 2 shows the CIRs over time by individual physicians. All colonoscopists had high CIRs at the outset, and this continued over the years. None of the physicians had a CIR of <96.6% in any year.

Figures 3 and 4 show the overall CIRs per physician and per year. CIR varied from 98.7% to 100% by physician and 99%–99.7% by year. The photography audit results are shown in Table 1 (Supplementary Digital Content 1, <http://links.lww.com/CTG/A235>). The overall cecal photo documentation rate (at least one convincing

picture of appendiceal orifice, terminal ileum or ileocecal valve) was 98.4% with 12 of the 16 physicians documenting cecal intubation with ≥ 1 relevant picture 100% of the time. Table 2 (Supplementary Digital Content 2, <http://links.lww.com/CTG/A236>) details the number of procedures needed to establish CIRs of >95%. This table is derived mathematically and not from the study data. Using the table, for physicians with an expected sample proportion of 98% CIR, only 84 procedures are required to establish a confidence interval with a margin of less than 3%. This would document with a 95% confidence that the said physician, taking into account the sample variation, is above a 95% threshold for cecal intubation.

Figure 5 shows the scatter plot comparing the number of procedures by the physician to their CIR. The performance of one physician with a much larger volume of colonoscopies than the other physicians in the study is not shown in this plot. No significant correlation was observed (spearman's rho: -0.196 , $P = 0.483$). Figures 1 and 2, Supplementary Digital Content 3 and 4, <http://links.lww.com/CTG/A231>, <http://links.lww.com/CTG/A232> show the results of including colonoscopies that were documented as no intent to intubate cecum. The overall CIR remained close to 99% and showed little variation over the 16-year period, similar to the previous results. Fifteen of the 16 doctors remained above 95% once they reached this threshold. One of the 16 had their CIRs drop substantially in the first 3 years of the practice at our center because of performing large numbers of colonoscopies for fecal microbiota transplant, which initially were not performed with intent to intubate the cecum.

Figures 3 and 4, Supplementary Digital Content 5 and 6, <http://links.lww.com/CTG/A233>, <http://links.lww.com/CTG/A234> show the CIR after inclusion of patients with inadequate bowel preparation. The 16 study physicians performed 2,967 of the 4,546 procedures designated with inadequate preparation. Sixty-seven procedures with inadequate preparation were also documented as having no intent for cecal intubation. The rates of identifying bowel preparation as inadequate varied from 1.9% to 12.6% among the 16 study physicians. The overall CIR in the 2,900 procedures performed by the 16 study physicians with inadequate preparation was

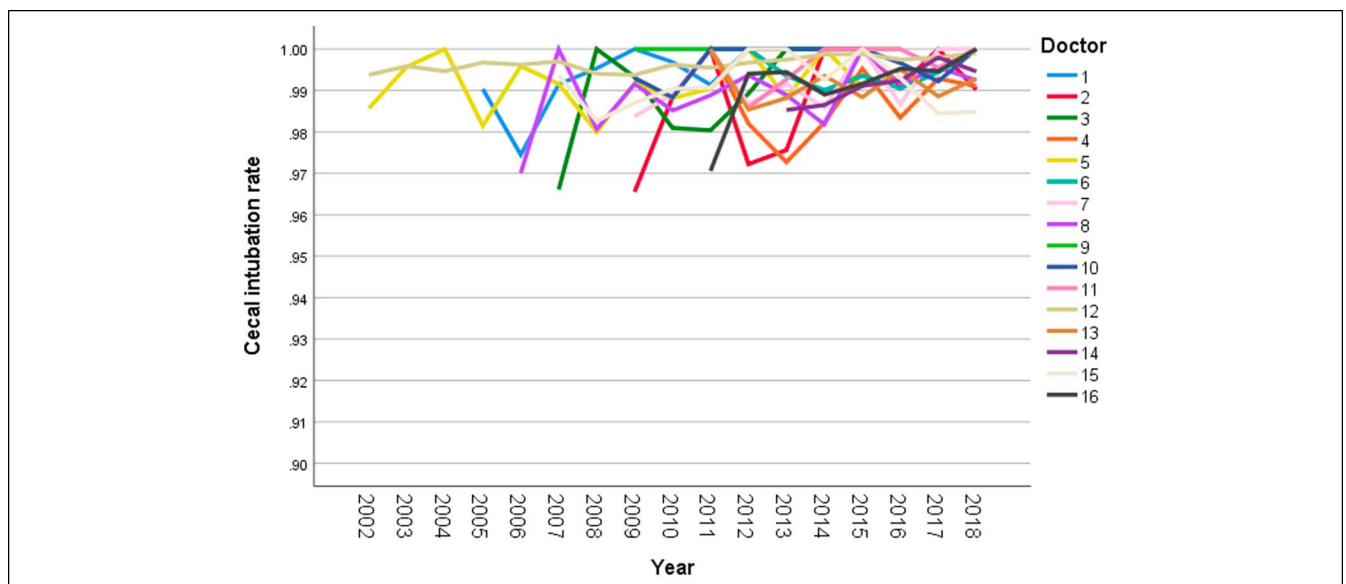


Figure 2. Line graph of cecal intubation rate for 16 colonoscopists each performing 50 procedures over the 6 consecutive years. Procedures with inadequate preparation and no intent to intubate the cecum are excluded.

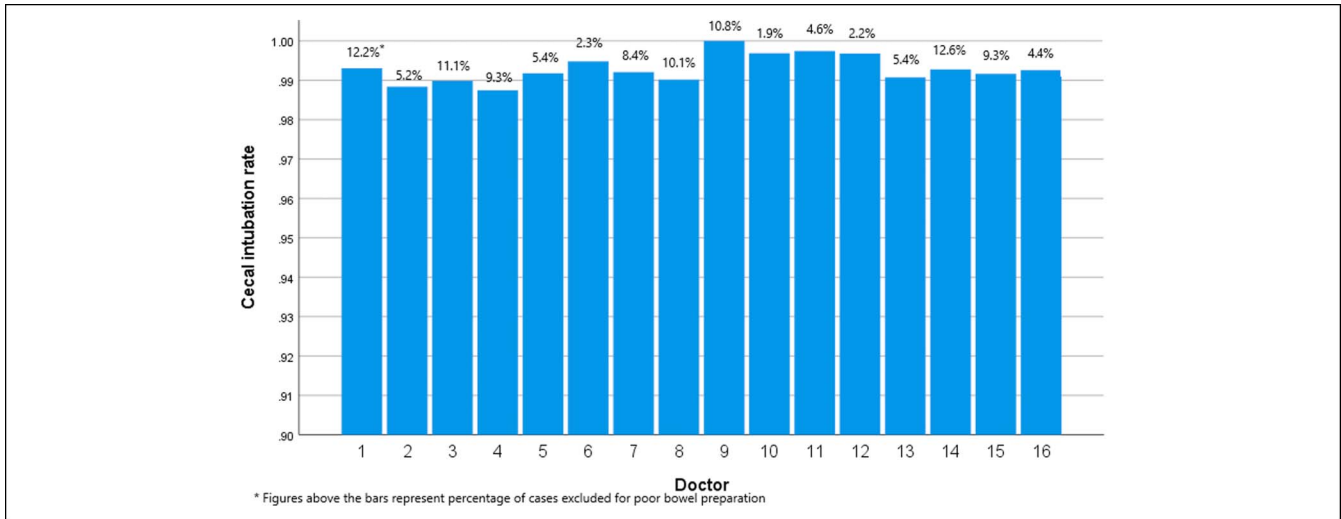


Figure 3. Cecal intubation rate by a colonoscopist over the full study period. Procedures with inadequate preparation and no intent to intubate the cecum are excluded. The fraction of all cases called inadequate by each physician is shown at the top of the columns.

72.3%. With inclusion of these procedures, the overall CIR for the 16 study physicians was still 98%.

DISCUSSION

We present data that cast doubt on the use of measuring cecal CIRs continuously over years for many gastroenterologists. Essentially, all gastroenterologists evaluated in our study had CIRs substantially above the recommended threshold for screening examinations and did not have declines in CIR over time. Furthermore, CIRs were high and remained high for most physicians, regardless of whether CIR was adjusted for factors such as poor preparation or lack of intent to intubate the cecum. Stopping or eliminating efforts to measure CIR over time could free resources to measure other quality parameters (1), including important colonoscopy outcomes such as polypectomy skill (11). Although not demonstrated here, we suspect that similar considerations apply to ADR.

Measurement of actual CIRs in electronic reporting systems requires minimum effort because the reports can be generated

electronically. We showed that for high performers, the electronic report is based on far more examinations than are needed to ensure with high confidence that the CIR is above the recommended minimum thresholds. This finding would have more significance in centers that do not use electronic report generating systems, where verification of CIR might require manual review of reports and data entry. Anecdotally, we continue to see many referrals to our center with previous colonoscopy reports that are dictated rather than created by electronic report generating systems. Thus, we suspect that simplifying the CIR review process and reducing the number of colonoscopies needed to review would be more impactful in the community compared with academic practice (see Table 2, Supplementary Digital Content 2, <http://links.lww.com/CTG/A236>).

Our data suggest that for examiners with measured CIRs above 99%, the measurement of CIR in all examinations performed requires many more examinations be reviewed than are actually needed to ensure that CIR is above a minimum threshold of 95%. For example, for the highest volume endoscopist in our group, reviewing

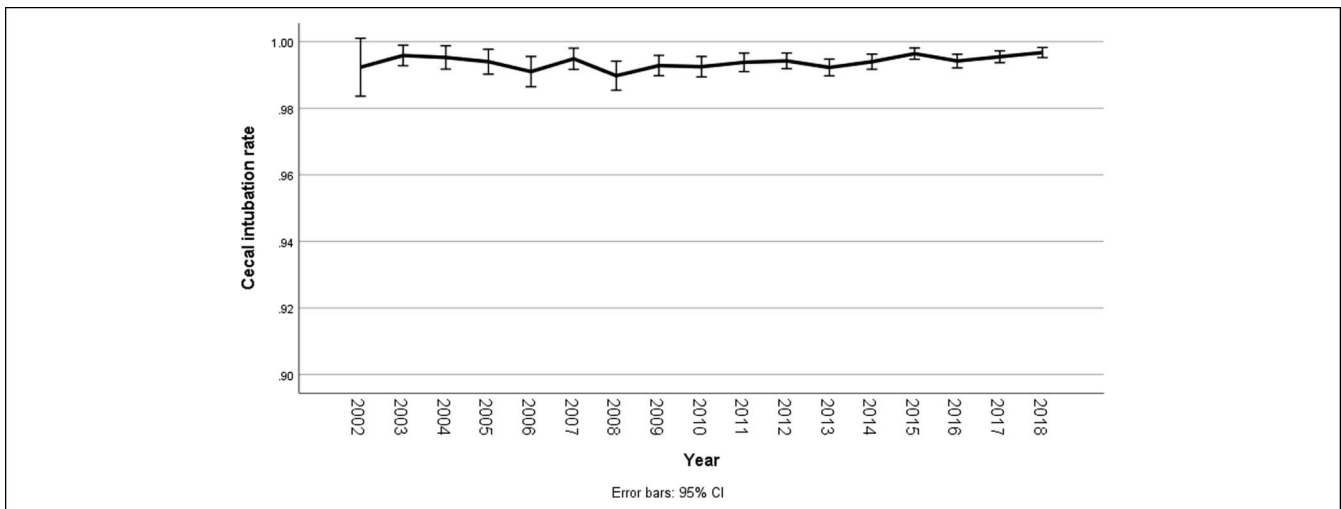


Figure 4. Overall cecal intubation rate for all colonoscopists, 2002–2018. Procedures with inadequate preparation and no intent to intubate the cecum are excluded.

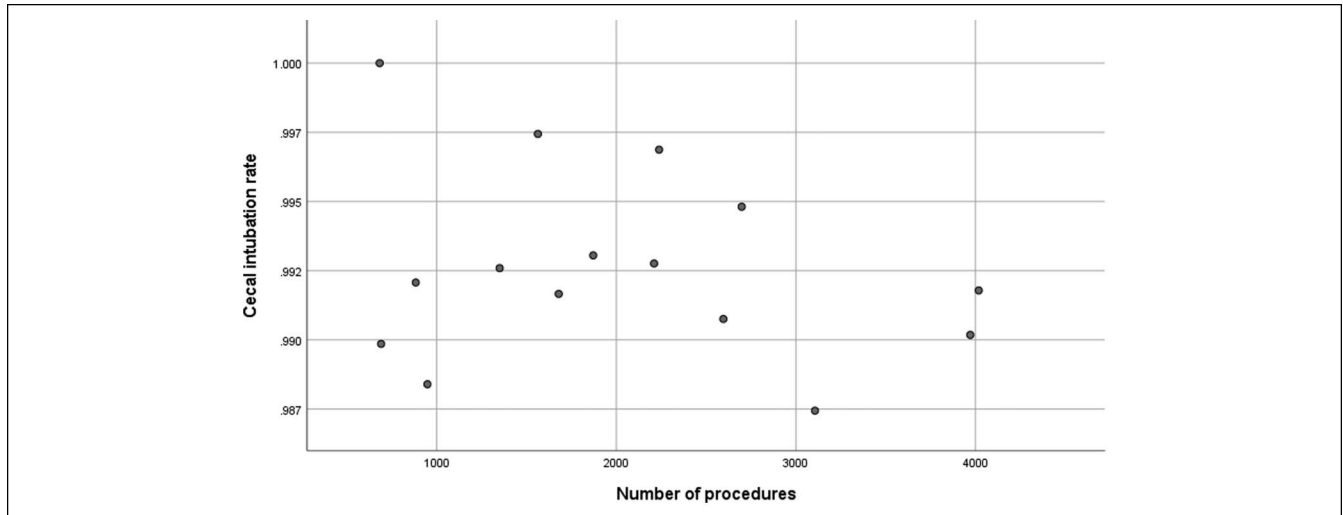


Figure 5. Scatter plot of number of procedures at our center vs the cecal intubation rate.

all procedures to document CIR entails 60 times more examinations than are needed to prove the CIR is $\geq 95\%$. If a 90% overall threshold is used for acceptable performance, the quality review process is even more redundant for high level performers. It seems reasonable to reduce the burden of measurement with the use of a smaller sample of cases to calculate CIR. For example, for doctors with CIRs $\geq 97\%$, it would be possible to intermittently review a sample of 280 colonoscopies to be confident that CIR remains above a minimum threshold of 95%. Audits of cecal photography are still generally necessary in quality programs. Our calculations show that the size of the audit can be substantially reduced for high level performers. In the future, artificial intelligence programs could also be used to verify adequate photo documentation of the cecum (12).

We did not observe an association between procedure volume at our center and CIRs. In previous studies, there was an association between procedure volume and CIR, particularly for doctors with ≤ 5 years of experience (9). We suspect that 2 factors are related to this finding in our study. First, all 16 doctors had high CIRs, which reduces the chance to observe an association between procedure volume and CIR. Second, although doctors may perform smaller numbers of colonoscopies in this study, these same endoscopists are often high volume endoscopists at other endoscopy units in our healthcare system. Thus, the procedure volume at the study center is not reflective of overall colonoscopy volumes and experience for these physicians.

An alternate or additional interpretation of these data is that target CIRs should be increased above 95%, perhaps to 97% or 98%. However, larger databases involving more colonoscopists from other centers should be examined to determine the appropriateness of changing recommended targets. Furthermore, our data still would support discontinuous measurement, given that CIRs showed no tendency in our study to decline with time.

The limitations of our study include that it was retrospective. In addition, the data are from a center recognized for endoscopy, which might limit generalizability. However, most of the endoscopists were general endoscopists and included individuals whose major emphasis is hepatology or motility. We acknowledge that it is appropriate to examine this issue in other centers. Strengths of the study include the large number of colonoscopies examined. Although the study was performed in an academic medical center, we

suspect that the results regarding the stability or improvement and lack of decline in CIRs over time are generalizable.

In summary, we have shown that for CIR measurement in a colonoscopy quality program, examiners with very high CIRs maintain these rates over time. Although measurement of CIR remains important (1), the use of repeated measurements over time in very high level performers could waste resources, depending on how easily the measurements are made. Freeing resources could allow for measurement of nonpriority indicators (1) or polypectomy skills (13,14).

CONFLICTS OF INTEREST

Guarantor of the article: Douglas K. Rex, MD.

Specific author contributions: All authors wrote the manuscript and approved the final version.

Financial support: This work was supported by a grant from Scott Schurz of Bloomington, Indiana, and his children to the Indiana University Foundation in the name of Douglas K. Rex.

Potential competing interests: D.K.R.: Consultant: Olympus Corporation, Boston Scientific, Medtronic, Aries Pharmaceutical, Braintree Laboratories; Research Support: EndoAid, Olympus Corporation, Medivators; Ownership: Satisfai Health. All other authors have no relevant relationships to disclose.

Study Highlights

WHAT IS KNOWN

- ✓ The cecal intubation rate is considered a priority quality indicator for colonoscopy.

WHAT IS NEW HERE

- ✓ In a single center study of 16 endoscopists, CIR was high and stayed high over at least 6 years. Colonoscopists with very high CIRs needed audit of smaller numbers of colonoscopies to verify CIR $> 95\%$.

TRANSLATIONAL IMPACT

- ✓ CIRs that are well above recommended thresholds could be monitored intermittently rather than continuously, which in some settings would free resources to measure other quality indicators.

REFERENCES

1. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. *Gastrointest Endosc* 2015;81:31–53.
2. Brenner H, Altenhofen L, Kretschmann J, et al. Trends in adenoma detection rates during the first 10 years of the German screening colonoscopy program. *Gastroenterology* 2015;149:356–66.e1.
3. Shaukat 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;7:1335–40.
4. Duloy AM, Kaltenbach TR, Keswani RN. Assessing colon polypectomy competency and its association with established quality metrics. *Gastrointest Endosc* 2018;87:635–44.
5. Rembacken B, Hassan C, Riemann JF, et al. Quality in screening colonoscopy: Position statement of the European Society of Gastrointestinal Endoscopy (ESGE). *Endoscopy* 2012;44:957–68.
6. Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. *Am J Gastroenterol* 2015;110:72–90.
7. Chilton A, Rutter M. BCSP Quality Assurance Guidelines for Colonoscopy. NHS Cancer Screening Programmes: Sheffield, United Kingdom, 2011.
8. Segnan N, Patnick J, von Karsa L. European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis, 1st ed. International Agency for Research on Cancer: Luxembourg, 2010.
9. Harewood GC. Relationship of colonoscopy completion rates and endoscopist features. *Dig Dis Sci* 2005;50:47–51.
10. Aslinia F, Uradomo L, Steele A, et al. Quality assessment of colonoscopic cecal intubation: An analysis of 6 years of continuous practice at a university hospital. *Am J Gastroenterol* 2006;101:721–31.
11. Pohl H, Srivastava A, Bensen SP, et al. Incomplete polyp resection during colonoscopy—results of the complete adenoma resection (CARE) study. *Gastroenterology* 2013;144:74–80 e1.
12. Karnes WE, Ninh A, Dao T, et al. Real-time identification of anatomic landmarks during colonoscopy using deep learning. *Gastrointest Endosc* 2018;87:Ab252.
13. Patel SG, Duloy A, Kaltenbach T, et al. Development and validation of a video-based cold snare polypectomy assessment tool (with videos). *Gastrointest Endosc* 2019;89:1222–30.e2.
14. Gupta S, Bassett P, Man R, et al. Validation of a novel method for assessing competency in polypectomy. *Gastrointest Endosc* 2012;75:568–75.

Open Access This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.