Title Page

Manuscript Number: CGH 18-00675

Title: Video-based Assessments of Colonoscopy Inspection Quality Correlate with Quality Metrics and Highlight Areas for Improvement

Short Title: Video-based Assessments of Colonoscopy Inspection Quality

Authors: Duloy, Anna, MD[1]; Yadlapati, Rena H, MD MS[2]; Benson, Mark, MD[3]; Gawron, Andrew J, MD PhD[4]; Kahi, Charles J, MD[5]; Kaltenbach, Tonya R, MD MS[6]; McClure, Jessica, MD[7]; Gregory, Dyanna L, BS[1]; Keswani, Rajesh N, MD MS[1]

1. Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine; Chicago, IL.

2. Department of Gastroenterology, University of Colorado; Aurora, CO.

3. Department of Gastroenterology and Hepatology, University of Wisconsin School of Medicine and Public Health; Madison, WI.

4. Division of Gastroenterology, Hepatology and Nutrition, University of Utah; Salt Lake City, UT.

5. Department of Gastroenterology, Indiana University School of Medicine, Richard L Roudebush VA Medical Center; Indianapolis, IN.

6. Department of Gastroenterology, University of California San Francisco; San Francisco, CA.

7. Department of Medicine, Northwestern University Feinberg School of Medicine; Chicago, IL.

Grant Support: none

Abbreviations: adenoma detection rate (ADR); colonoscopy inspection quality (CIQ); colorectal cancer (CRC); interquartile range (IQR); serrated polyp detection rate (SDR); withdrawal time (WT)

Correspondence: Anna Duloy; e-mail: anna.duloy@northwestern.edu; phone: 615-414-4859; address: 676 North Saint Clair, Suite 1400, Chicago, IL 60611

Disclosures: Anna Duloy, Mark Benson, Andrew Gawron, Charles Kahi, Dyanna Gregory, and Jessica McClure have no disclosures. Rena Yadlapati is supported by NIH T32DK101363 grant. Tonya Kaltenbach is a consultant for Olympus America. Rajesh Keswani is a consultant for Boston Scientific and Medtronic.

Writing Assistance: none

Author Contributions: Anna Duloy contributed to the conception and design of the study; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; and statistical analysis. Rena Yadlapati contributed to the analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content. Mark Benson contributed to the analysis and interpretation of data. Andrew Gawron contributed to the analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important

intellectual content. Charles Kahi contributed to the analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content. Jessica McClure contributed to the acquisition of data. Dyanna Gregory contributed to the acquisition of data and analysis and interpretation of data. Tonya Kaltenbach contributed to the analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content. Rajesh N. Keswani contributed to the study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis; and study supervision. All authors have approved the final draft submitted.

Abstract

Background & Aims: Adenoma detection rate (ADR) and serrated polyp detection rate (SDR) vary significantly among colonoscopists. Colonoscopy inspection quality (CIQ) is the quality with which a colonoscopist inspects for polyps and may explain some of this variation. We aimed to determine the relationship between CIQ and historical ADRs and SDRs in a cohort of colonoscopists and assess whether there is variation in CIQ components (fold examination, cleaning, and luminal distension) among colonoscopists with similar ADRs and SDRs.

Methods: We conducted a prospective observational study to assess CIQ among 17 high-volume colonoscopists at an academic medical center. Over 6 weeks, we video-recorded >28 colonoscopies per colonoscopist and randomly selected 7 colonoscopies per colonoscopist for evaluation. Six raters graded CIQ using an established scale, with a maximum whole colon score of 75.

Results: We evaluated 119 colonoscopies. The median whole-colon CIQ score was 50.1/75. Whole-colon CIQ score (r=0.71; *P*<.01) and component scores (fold examination r=0.74; cleaning r=0.67; distension r=0.77; all *P*<.01) correlated with ADR. Proximal colon CIQ score (r=0.67; *P*<.01) and component scores (fold examination r=0.71; cleaning r=0.62; distension r=0.65; all *P*<.05) correlated with SDR. CIQ component scores differed significantly between colonoscopists with similar ADRs and SDRs for most of the CIQ skills.

Conclusion: In a prospective observational study, we found CIQ and CIQ components to correlate with ADR and SDR. Colonoscopists with similar ADRs and SDRs differ in their performance of the 3 CIQ components—specific, actionable feedback might improve colonoscopy technique.

KEY WORDS: quality improvement, endoscopy, early detection, colon cancer prevention

INTRODUCTION:

The effectiveness of screening colonoscopy in the prevention of colorectal cancer (CRC) relies upon the quality of its performance, specifically the detection of neoplastic colon polyps. The adenoma detection rate (ADR) is regarded as the primary indicator for the quality of mucosal inspection during colonoscopy, with an inverse association between colonoscopist ADR and risk of interval CRC in large cohort studies.[1,2] In addition to adenomas, colonoscopists must identify serrated polyps, which are a significant contributor to CRC and account for a disproportionate fraction of interval CRCs. Due to their flat morphology and location in the proximal colon, serrated polyps may be more difficult to identify during colonoscopy than conventional adenomas.[3] Importantly, both serrated polyp detection rate (SDR) and ADR vary significantly among colonoscopists.[4-9]

The quality with which a colonoscopist inspects the colon for polyps may explain some of the observed variations in ADR and SDR. A scale to evaluate colonoscopy inspection quality (CIQ) was developed by Rex[10] and has previously been shown to correlate with ADR.[10,11] The CIQ scale assesses performance on three complementary skills: fold examination, luminal distension, and mucosal cleaning. The smaller size of the prior studies limited determination of which CIQ factors are most associated with ADR. Further, the association between CIQ and SDR has not been studied. Evaluating how colonoscopists perform on these specific skills could allow for targeted feedback and individualized improvement strategies.

The primary aim of our study was to determine the relationship between CIQ and historical ADR and SDR among a large cohort of colonoscopists with varying baseline ADRs and SDRs. We

hypothesized that superior fold examination would correlate with increasing colonoscopist ADR and superior fold examination in the proximal colon would correlate with increasing colonoscopist SDR. Our secondary aim was to determine if variation in individual CIQ components (fold examination, cleaning, and luminal distension) exists among colonoscopists with similar ADRs/SDRs.

METHODS:

Study Design & Setting:

We conducted a prospective observational study of colonoscopists performing screening and surveillance colonoscopy at a single urban academic medical center from 10/3/2016 to 11/11/2016. The Northwestern University Institutional Review Board approved the study (IRB #: STU00203769, approval date 9/8/2016). Colonoscopists included in the study provided written informed consent.

We recruited colonoscopists who had performed 100 or more annual screening colonoscopies in the two years preceding the study onset. Over a six-week period (10/3/2016 – 11/11/2016), study investigators prospectively recorded at least 28 de-identified screening or surveillance colonoscopies performed by each colonoscopist. We excluded colonoscopies performed for diagnostic indications, inflammatory bowel disease, or a personal history of a polyposis syndrome or cancer. We also excluded colonoscopies with a Boston Bowel Preparation Score less than six and colonoscopies performed with a trainee.

Video-recorders were set-up by a single study investigator not participating in the colonoscopy rating process. Video recordings were obtained utilizing a portable high-definition digital video recorder (Sony HVO-500MD) attached to the digital endoscope processor. Patient and physician identifiers were removed from the TV monitor prior to the start of the recordings. The colonoscopists were aware of the recorders, but not of when they were specifically being recorded.

Seven videos per colonoscopist were randomly selected using a random number generator. CIQ was evaluated by six U.S. gastroenterologists (RY, MB, AG, CK, TK, RK) with previous experience in colonoscopy quality ("colonoscopy raters").

Colonoscopy Inspection Quality (CIQ):

CIQ was assessed using a scale developed by Rex[10] and adapted by Lee et al.[11] To assess CIQ, colonoscopy raters evaluated the entire colonoscopy withdrawal, assigning segmental scores from 0 to 5 based on the adequacy of three components: fold examination, cleaning, and luminal distension. We defined the scores as: 0=very poor (not looking behind any folds, "straight pull-back" technique; no attempt to clean any stool/pools of liquid; no colonic distension or spasm present), 1=poor, 2=fair, 3=good, 4=very good, 5=excellent (looking behind all folds; stool/pools of liquid removed; full colonic distension to allow for ideal mucosal visualization). Colon segments where there were no pools of liquid and therefore did not require any cleaning received a cleaning score of 5. A total of five colon segments were scored (cecum/appendiceal orifice/ileocecal valve; ascending colon; transverse colon; descending colon; and sigmoid/rectum) for each CIQ component, for a maximum score of 75 in the whole colon [Table

1]. The maximum score in the proximal colon (defined as cecum through transverse colon) was 45 and the maximum score in the left colon (descending colon to rectum) was 30. In studies where cecal retroflexion was performed, the raters specifically evaluated if there was improved fold examination in the retroflexed view; however, if the endoscopist simply retroflexed without any added benefit, retroflexion did not result in an improvement in the fold examination score. Notably, endoscopists who did not perform retroflexion, were not penalized.

Colonoscopy raters were also asked to assess: 1) the number of complete evaluations of the right colon (defined as the number of complete passes from the cecum to the hepatic flexure in forward or retroflexed view); 2) whether cecal retroflexion was performed; 3) a qualitative, binary, assessment of whether inspection was "adequate" in each of the five colon segments; and 4) the time the colonoscope last reached the hepatic flexure and splenic flexure to calculate segmental withdrawal times.

To standardize the review process, all six raters initially graded four videos and inter-rater score variation was discussed between the colonoscopy raters to determine sources of variation and agree upon scoring criteria and individual skill meaning. Five raters graded one video per colonoscopist and one rater graded two videos per colonoscopist. Raters were blinded to the colonoscopist and reviewed the videos independently and in random order.

Study Outcomes:

The primary study outcomes were CIQ and colonoscopist historical ADR and SDR.

Data Sources and Measurement:

ADR, SDR, and withdrawal time (WT) were calculated using twelve-month historical data (8/1/2015 – 7/31/2016) of screening colonoscopies performed by each colonoscopist. Data was obtained from our institution's Enterprise Data Warehouse, a single, integrated database of clinical and research information from all patients receiving treatment through Northwestern University healthcare affiliates.

A screening colonoscopy was defined as a colonoscopy in a patient aged 50 to 75, with an indication of detecting colorectal neoplasia. Patients with a prior history of colon adenomas/serrated polyps, or a colonoscopy performed to evaluate signs or symptoms of gastrointestinal pathology including occult blood loss, anemia, abdominal pain, or rectal bleeding were excluded. ADR was defined as the proportion of screening colonoscopies with $\geq$1 adenoma and SDR was defined as the proportion of screening colonoscopies with $\geq$1 sessile serrated polyp or traditional serrated adenoma; hyperplastic polyps were not included in the SDR. Historical withdrawal time was defined as the time spent withdrawing the colonoscope (inspecting for polyps) in screening colonoscopies where no pathology was obtained (i.e., no polyps found, and no biopsies taken).

Study withdrawal times were manually calculated from the study videos. Total WT was defined as the time from the identification of cecal landmarks to scope removal from the rectum, excluding any time spent in the ileum and time spent performing polypectomy and/or biopsy. WT to the hepatic flexure was defined as the time from the cecum to the hepatic flexure minus

time spent performing polypectomy/biopsy and WT to the splenic flexure was defined as the time from the cecum to the splenic flexure minus time spent performing polypectomy/biopsy.

Statistical Analysis:

A complete case analysis was performed, we did not anticipate any missing data, and all analyses were planned *a priori*. CIQ scores per colonoscopist were averaged and the median average reported ("median"). The primary analyses examined the relationship between colonoscopist CIQ and historical ADR, historical SDR, and WT (study and historical) using Spearman correlation. The secondary analyses examined variability in CIQ component scores among colonoscopists in the same ADR and SDR tertiles. We used one-way ANOVA to assess variation in CIQ component scores within each tertile and examined the relationship between component scores using repeated measures ANOVA. To examine inter-rater reliability, we performed a sensitivity analysis assessing intraclass correlations for consistency and absolute agreement based on a two-way random effects model. P-values less than .05 were considered statistically significant. Sample size calculations were based on both primary and secondary analyses. With seven videos for each of the 17 colonoscopists, we had a minimum of 80% power to detect significant correlations of 0.64 and CIQ mean component score differences of at least 4 by colonoscopist, assuming a standard deviation in component score of 2 and a Type I error rate of 5%. SAS 9.4 (Cary, NC) was used for all main statistical analyses. IBM SPSS Statistics 24 was used for inter-rater reliability analysis.

RESULTS:

Seventeen colonoscopists (16 gastroenterologists and 1 colorectal surgeon) met inclusion criteria and provided informed consent.

Historical ADR, SDR, and WT:

The 17 colonoscopists performed a median of 424 screening colonoscopies (range 108-775) in the 12 months preceding study onset. The median historical ADR was 38% (Interquartile Range [IQR] 31%-44%) and median SDR was 10% (IQR 8%-13%). Median historical WT was 11.1 minutes (IQR 8.7 - 14.0 minutes).

CIQ and Study WT:

A total of 504 videos were recorded during the study period and 119 videos were graded. Median whole colon CIQ score (maximum 75) was 50.1 (IQR 44.3 - 57.7). The median proximal colon CIQ score (maximum 45) was 30.1 (IQR 27.3 - 36.1) and the median left colon CIQ score (maximum 30) was 19.6 (IQR 17.1 - 21.6). Among the 3 CIQ component scores (maximum 25 each), the median fold examination score was significantly lower (14.9; IQR 11.7 - 16.3) than the cleaning (18.6; IQR 16.9 - 21.4, P<.01) and distension (17.1; IQR 15.9 - 20.7, P<.01) scores.

Cecal retroflexion was performed in 32% of colonoscopies. More than one complete examination of the right colon was performed in 27% of colonoscopies. The median total study WT was 12.6 minutes (IQR 10 - 14.7 minutes) with significantly more time spent in the proximal colon (7.4 minutes) compared to the left colon (4.9 minutes, P<.01).

Inter-Rater Reliability:

To test inter-rater reliability of the whole colon CIQ score, six raters each rated the same six randomly selected colonoscopies performed by six different colonoscopists. Intraclass correlations for both consistency and absolute agreement were calculated based on a two-way random effects model. The intraclass correlation for consistency was 0.94 (95% CI 0.82, 0.99), while the intraclass correlation for absolute agreement was 0.73 (95% CI 0.33, 0.95).

Primary Analysis - Relationship between Colonoscopist CIQ and ADR/SDR/WT:

Median whole colon CIQ score (r=0.71, P<.01) and all 3 CIQ component scores (fold examination r=0.74; cleaning r=0.67; distension r=0.77; all P<.01) significantly correlated with ADR [Table 2]. The number of segments with suboptimal inspection, as assessed by the raters, negatively correlated with ADR (r=-0.70, P<.01).

Median whole colon CIQ score (r=0.62, P<.01) and all 3 CIQ component scores (fold examination r=0.67; cleaning r=0.54; distension r=0.63; all P<.05) significantly correlated with SDR [Table 3]. Because serrated polyps are most commonly found in the proximal colon, we evaluated the association between proximal colon CIQ scores and SDR. Proximal colon CIQ score (r=0.67, P<.01) significantly correlated with SDR, whereas performance of cecal retroflexion (r=0.12, P=.65) and number of complete examinations of the right colon (r=0.15, P=.58) did not. Proximal colon fold examination was most highly correlated with SDR (r=0.71, P<.01).

Only one colonoscopist had an ADR of less than 25% (12-month historical ADR of 18%); this colonoscopist also had the lowest SDR of the included colonoscopists (4%). The CIQ score for

this colonoscopist was the lowest (25.4) of all colonoscopists and was 11 points lower than any other colonoscopist.

Historical WT significantly correlated with CIQ (r=0.66, P<.01; [Figure 1]), ADR (r=0.70, P<.01), and SDR (r=0.54, P=.02). WT to the splenic flexure similarly correlated with SDR (r=0.52, P=.03). Despite the correlation between CIQ and WT, there were three high WT colonoscopists (>11.2-minute historical median WT) with whole colon CIQ scores below the cohort median. Mean study WT per colonoscopist was significantly longer than mean historical WT (+1.1 minutes, P=.03).

Secondary Analysis - Variation in CIQ Component Scores among Colonoscopists:

To assess for variation in CIQ component scores between colonoscopists with similar ADRs and SDRs, colonoscopists were divided into tertiles based on ADR (tertile 1: ADR 18-31%; tertile 2: ADR 33-39%; tertile 3: ADR 41-57%) and SDR (tertile 1: SDR 5-7%; tertile 2: 8-10%; tertile 3: 12-20%). Mean CIQ scores (whole colon for ADR and proximal colon for SDR) for fold examination, cleaning, and distension among all the colonoscopists were calculated and each colonoscopists' individual performance in fold examination, cleaning, and distension were plotted relative to the mean [Figures 2 and 3]. CIQ component scores were significantly different between colonoscopists within the same ADR tertile (all P<.05) for most of the groups. There were significant differences (all P<.05) for proximal colon component scores in all the SDR tertile groups.

DISCUSSION:

We performed a prospective observational study of 17 colonoscopists to assess metrics of colonoscopy quality using a previously validated scale.[10,11] This study demonstrates that colonoscopy inspection quality and its components (fold examination, cleaning, and distension) strongly correlate with ADR and SDR. In addition, colonoscopists with similar ADRs/SDRs vary significantly in performance of CIQ components. These results suggest that assessment of CIQ is a valid and reliable metric of colonoscopy quality for adenoma and serrated polyp detection and can be calculated from a relatively low number of colonoscopies. Moreover, CIQ scores highlight colonoscopists' strengths and weaknesses, which vary by colonoscopist, providing actionable targets for practice improvement.

The strong correlation between CIQ and ADR in this study is consistent with findings from prior validation studies by Rex[10] and Lee et al.[11] In the study done by Rex, video recordings of colonoscopy withdrawals performed by two colonoscopists with different adenoma miss rates (17% vs. 46%) were graded by four experts. The colonoscopist with the lower miss rate had superior CIQ scores. Similarly, Lee et al graded video-recordings of colonoscopy withdrawals performed by 11 colonoscopists who were divided into low, moderate, and high ADR groups. They found that colonoscopists with high or moderate ADRs had superior mean CIQ scores compared with colonoscopists in the low ADR group. In the largest group of colonoscopists to date, we similarly found that CIQ strongly correlates with ADR and is most strongly correlated with fold examination and luminal distension, suggesting that these practices are integral to colonoscopy quality.

Consistent with recent studies,[3-4,9] the range of serrated polyp detection rates among the colonoscopists in our study was highly variable, ranging from 5-20%. The source of this variation has not been well studied and no prior study has evaluated the relationship between CIQ and SDR. Thus, we aimed to evaluate the association between colonoscopy technique, as measured by CIQ, and variations in SDR. Sessile serrated polyps are typically flat, located in the proximal colon, and may be "buried" in between deep folds.[3] As hypothesized, SDR was most strongly correlated with the proximal colon fold examination score. Interestingly, while withdrawal time to the hepatic and splenic flexures was strongly associated with SDR, we did not find an independent correlation between the number of complete examinations of the ascending colon and SDR. These results suggest that performing quick additional withdrawals without optimal inspection technique do not increase colonoscopy quality.

We expected to see a significant correlation between cecal retroflexion and SDR but did not find one. We found that the majority of colonoscopists were quickly retroflexing in the cecum, without performing a complete second examination in the retroflexed view. Cecal retroflexion as a maneuver, without standardized training and assessment, is unlikely to provide added benefit. Whether a second examination – in forward or retroflexed view – increases SDR requires further study.

The mandate for a minimum colonoscopy withdrawal is largely based on the assumption that WT serves as a proxy for the quality of colonoscopy inspection.[12,13] However, it is unlikely that increasing WT alone will improve the quality of inspection and previous studies regarding WT have shown conflicting results.[14,15] We found that WT was significantly correlated with both

ADR (r=0.70, P<.01) and SDR (r=0.54, P=.02). We were unable to determine whether CIQ and WT were independently associated with improving ADR and SDR, because of the very strong correlation between CIQ and WT (r=0.66, P=<.01). We did, however, find a wide variation in CIQ scores for colonoscopists with similar WTs. Furthermore, some colonoscopists with longer WTs had CIQ scores which were lower than study mean. Similarly, while Rex[10] found significant differences in WT between the colonoscopist with the high and low adenoma miss rates (6.7 minutes vs. 8.9 minutes, P=.02), Lee et al[11] did not find differences in mean WT between the colonoscopists in the low, moderate, and high ADR groups. In the latter study, this lack of variation was attributed to the Hawthorne effect, with colonoscopists with lower ADRs "playing to the clock" and slowing down their withdrawal speeds to meet a 6-minute goal. In summary, the data suggests that performing a high-quality colonoscopy generally takes additional inspection time but that a longer WT does not automatically ensure a high-quality examination.

There is increasing interest in providing feedback to colonoscopists to drive quality improvement efforts.[16,17] However, while colonoscopist ADR is a critical metric of colonoscopy quality, it does not highlight the specific skills in need of improvement. Similarly, SDR is not routinely measured, does not have validated benchmarks, and does not provide targeted strategies for improvement. In our study, we found that there was significant heterogeneity in CIQ scores despite stratification by ADR and SDR performance level. In other words, despite similar ADRs and SDRs, colonoscopists generally had significantly different skills related to fold examination, distention and cleaning. Individualized feedback regarding specific examination skills could allow for targeted improvement strategies. As CIQ can be calculated using a small number of

colonoscopies (unlike ADR which has limited use among low-volume colonoscopists[18]), does not require advanced data analytics, correlates significantly with ADR and SDR, and clearly identifies underperforming colonoscopists, it is conceivable that CIQ may complement standard quality metrics for some colonoscopists. Future studies should explore feasible mechanisms to rate CIQ (i.e., having colonoscopists grade each other) and the effect of targeted CIQ feedback on endoscopic performance.

While CIQ can explain some of the variation in ADR and SDR, our results highlight that there are additional factors playing a role. For example, we identified some colonoscopists with higher ADR/SDRs with lower CIQ scores. Additionally, we found significant variation in CIQ component scores among endoscopists in the same ADR/SDR tertiles. We suspect that some of these differences are related to differences in withdrawal time. However, it is likely that additional factors that are not accounted for by CIQ, such as visual acuity, visual gaze, and education/recognition of nonpolypoid lesions are of importance.[19-22]

This study has several strengths. This is the largest study to evaluate colonoscopy inspection quality and the first to examine the relationship between CIQ and SDR. There are also several limitations. First, there is a risk of bias from the Hawthorne effect as several of the colonoscopists were aware they were being recorded and may have modified their behaviors. However, despite this, we have shown that the high correlation between CIQ and historical ADR and SDR remains, suggesting a limited extent of behavior modification and/or ability to modify technique. Second, we used videos of both screening and surveillance colonoscopies to prospectively grade CIQ. However, we still found that CIQ correlated significantly with

screening colonoscopy ADR and SDR. Third, reviewing colonoscopy withdrawals may be burdensome. However, our expert raters had a high inter-rater reliability; this should facilitate multiple raters reviewing fewer videos and result in less individual rater burden. Fourth, our study cohort had low numbers of under-performing colonoscopists. However, the colonoscopist with the lowest ADR (ADR 18%) had the lowest CIQ of all endoscopists by an 11-point margin, suggesting that CIQ is sufficiently robust at delineating low-performing endoscopists. Finally, we acknowledge the greater difficulty in accurately identifying serrated polyps compared to adenomas and the associated limitation of SDR as a quality metric. Our SDR calculation was based on the presence of sessile serrated polyps and traditional serrated adenomas and did not include proximal hyperplastic polyps because this data was not available. However, we limited the SDR calculation to the 12 months preceding study onset, encompassing a time when our pathologists have a heightened awareness of the importance of correctly identifying serrated lesions.

In conclusion, we have demonstrated that a blinded assessment of colonoscopist CIQ strongly correlates with established metrics of colonoscopy quality. Thus, measuring CIQ using video-recordings may facilitate identification of low performing colonoscopists in centers that cannot easily calculate ADR and SDR. Furthermore, assessing CIQ might also provide individual colonoscopists with more tailored feedback to drive quality improvement efforts, supporting our goal to improve the quality of colonoscopy and reduce the risk of CRC.

FIGURE LEGENDS:

Figure 1. Colonoscopist historical WT is significantly associated with CIQ (r=0.66; *P*<.01). However, 3 colonoscopists with WTs above the mean have CIQ scores below the mean.

Figure 2. Individual CIQ component scores differ significantly among colonoscopists with similar ADRs. **P*<.05

Figure 3. Individual CIQ component scores differ significantly among colonoscopists with similar SDRs. **P*<.05

REFERENCES:

1. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;370:1298-306.
2. Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;362:1795-803.
3. Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 2012;107:1315-29; quiz 4, 30.
4. Kahi CJ, Hewett DG, Norton DL, Eckert GJ, Rex DK. Prevalence and variable detection of proximal colon serrated polyps during screening colonoscopy. Clin Gastroenterol Hepatol 2011;9:42-6.
5. Chen SC, Rex DK. Colonoscopist can be more powerful than age and male gender in predicting adenoma detection at colonoscopy. Am J Gastroenterol 2007;102:856-61.
6. Imperiale TF, Glowinski EA, Juliar BE, Azzouz F, Ransohoff DF. Variation in polyp detection rates at screening colonoscopy. Gastrointest Endosc 2009;69:1288-95.
7. Simmons DT, Harewood GC, Baron TH, et al. Impact of colonoscopist withdrawal speed on polyp yield: implications for optimal colonoscopy withdrawal time. Aliment Pharmacol Ther 2006;24:965-71.
8. Kahi CJ, Anderson JC, Rex DK. Screening and surveillance for colorectal cancer: state of the art. Gastrointest Endosc 2013;77:335-50.
9. Hetzel JT, Huang CS, Coukos JA, et al. Variation in the detection of serrated polyps in an average risk colorectal cancer screening cohort. Am J Gastroenterol 2010;105:2656-64.
10. Rex DK. Colonoscopic withdrawal technique is associated with adenoma miss rates. Gastrointest Endosc 2000;51:33-6.
11. Lee RH, Tang RS, Muthusamy VR, et al. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). Gastrointest Endosc 2011;74:128-34.
12. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006;355:2533-41.
13. Rex DK, Petrini JL, Baron TH, et al. Quality indicators for colonoscopy. Am J Gastroenterol 2006;101:873-85.
14. Barclay RL, Vicari JJ, Greenlaw RL. Effect of a time-dependent colonoscopic withdrawal protocol on adenoma detection during screening colonoscopy. Clin Gastroenterol Hepatol 2008;6:1091-8.
15. Sawhney MS, Cury MS, Neeman N, et al. Effect of institution-wide policy of colonoscopy withdrawal time > or = 7 minutes on polyp detection. Gastroenterology 2008;135:1892-8.
16. Keswani RN, Yadlapati R, Gleason KM, et al. Physician report cards and implementing standards of practice are both significantly associated with improved screening colonoscopy quality. Am J Gastroenterol 2015;110:1134-9.
17. Kahi CJ, Ballard D, Shah AS, Mears R, Johnson CS. Impact of a quarterly report card on colonoscopy quality measures. Gastrointest Endosc 2013;77:925-31.
18. Do A, Weinberg J, Kakkar A, Jacobson BC. Gastrointest Endosc 2013;77:376-80.
19. Lami M, Singh H, Dilley JH, et al. Gaze patterns hold key to unlocking successful search strategies and increasing polyp detection rate in colonoscopy. Endoscopy 2018.

20. Almansa C, Shahid MW, Heckman MG, Preissler S, Wallace MB. Association between visual gaze patterns and adenoma detection rate during colonoscopy: a preliminary investigation. Am J Gastroenterol 2011;106:1070-4.
21. Coe SG, Crook JE, Diehl NN, Wallace MB. An Endoscopic Quality Improvement Program Improves Detection of Colorectal Adenomas. Am J Gastroenterol 2013;108:219-26.
22. Ussui V, Coe S, Rizk C, Crook JE, Diehl NN, Wallace MB. Stability of Increased Adenoma Detection at Colonoscopy. Follow-up of an Endoscopic Quality Improvement Program-EQUIP-II. Am J Gastroenterol 2015;110:489-96.

TABLES:

Table 1. Colonoscopy Inspection Quality (CIQ) Scores by Colon Segment

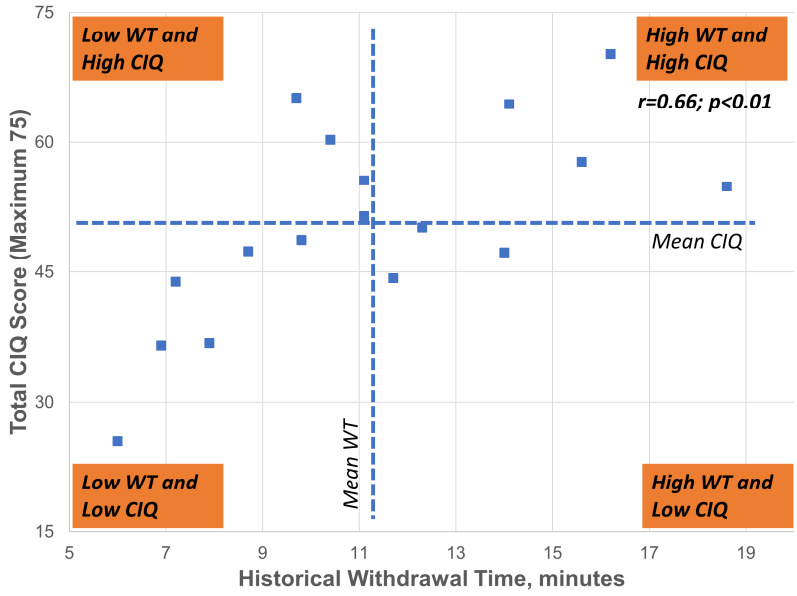
	Fold Examination Score	Cleaning Score	Luminal Distension Score	Colon Segment CIQ Score
Cecum, Appendiceal Orifice & Ileocecal Valve	0 – 5	0 – 5	0 – 5	0 – 15
Ascending Colon	0 – 5	0 – 5	0 – 5	0 – 15
Transverse Colon	0 – 5	0 – 5	0 – 5	0 – 15
Descending Colon	0 – 5	0 – 5	0 – 5	0 – 15
Sigmoid & Rectum	0 – 5	0 – 5	0 – 5	0 – 15
Whole Colon CIQ Score	0 – 25	0 – 25	0 – 25	0 – 75

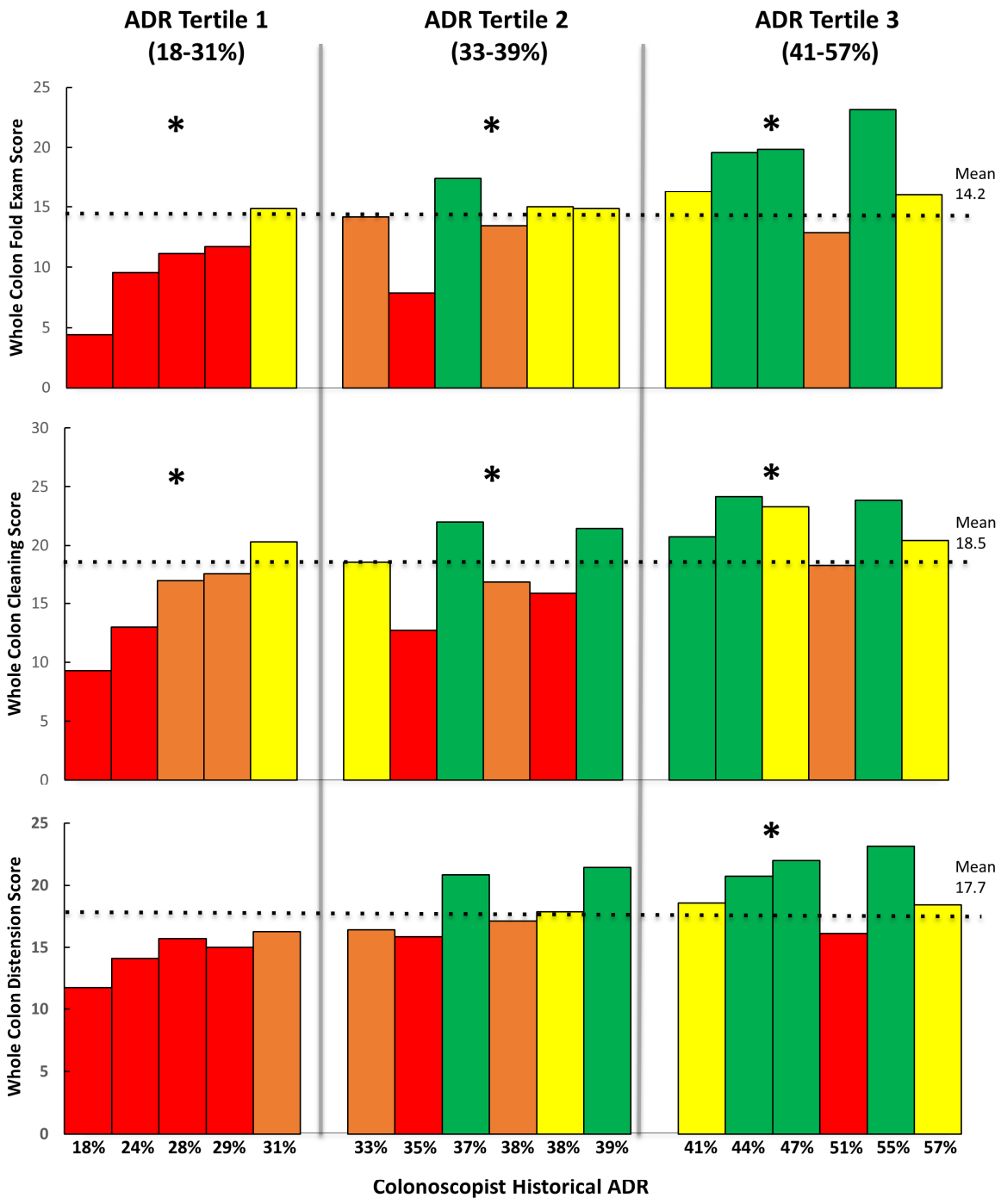
Table 2. Correlations between Colonoscopy Inspection Quality (CIQ) and Adenoma Detection Rate, Withdrawal Time, and Withdrawal Characteristics

Factor	**Spearman Correlation**	**95% Confidence Limits**	***P*-value**
CIQ Scores			
Whole Colon CIQ Score	0.71	(0.32, 0.88)	<0.01
Whole Colon CIQ Score, Fold Exam	0.74	(0.38, 0.89)	<0.01
Whole Colon CIQ Score, Cleaning	0.67	(0.26, 0.86)	<0.01
Whole Colon CIQ Score, Distention	0.77	(0.44, 0.91)	<0.01
Study Withdrawal Times			
Total Withdrawal Time	0.64	(0.22, 0.85)	<0.01
Withdrawal Time to Splenic Flexure	0.61	(0.17, 0.84)	0.01
Additional Withdrawal Characteristics			
Number of Complete Examinations of Right Colon	0.44	(-0.06, 0.76)	0.08
Cecal Retroflexion Performed	0.19	(-0.32, 0.61)	0.46
Number of Segments Requiring Further Inspection			
Cecum to Hepatic Flexure	-0.71	(-0.88, -0.33)	<0.01
Cecum to Splenic Flexure	-0.72	(-0.89, -0.34)	<0.01
Cecum to Rectum	-0.70	(-0.88, -0.31)	<0.01

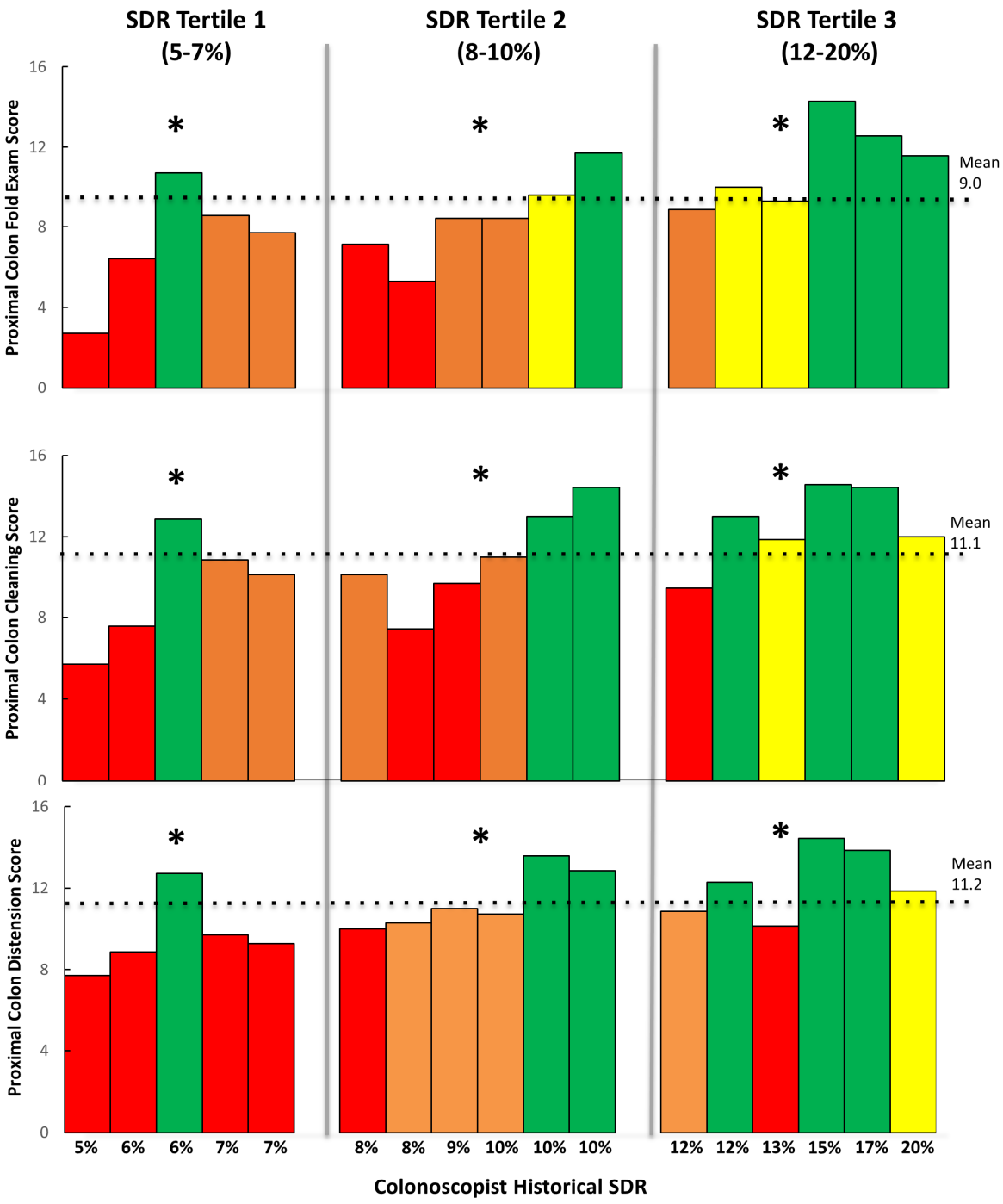
Table 3. Correlations between Colonoscopy Inspection Quality (CIQ) and Serrated Polyp Detection Rate, Withdrawal Time, and Withdrawal Characteristics

Factor	**Spearman Correlation**	**95% Confidence Limits**	***P*-value**
CIQ Scores			
Whole Colon CIQ Score	0.62	(0.18, 0.84)	0.01
Whole Colon CIQ Score, Fold Exam	0.67	(0.26, 0.87)	<0.01
Whole Colon CIQ Score, Cleaning	0.54	(0.06, 0.80)	0.03
Whole Colon CIQ Score, Distention	0.63	(0.19, 0.85)	0.01
Proximal Colon CIQ Score	0.67	(0.26, 0.87)	<0.01
Proximal Colon CIQ Score, Fold Exam	0.71	(0.34, 0.88)	<0.01
Proximal Colon CIQ Score, Cleaning	0.62	(0.18, 0.84)	0.01
Proximal Colon CIQ Score, Distention	0.65	(0.23, 0.86)	<0.01
Study Withdrawal Times			
Total Withdrawal Time	0.55	(0.08, 0.81)	0.02
Withdrawal Time to Splenic Flexure	0.52	(0.04, 0.80)	0.03
Additional Withdrawal Characteristics			
Number of Complete Examinations of Right Colon	0.15	(-0.36, 0.58)	0.58
Cecal Retroflexion Performed	0.12	(-0.39, 0.56)	0.65
Number of Segments Requiring Further Inspection			
Cecum to Hepatic Flexure	-0.61	(-0.84, -0.17)	0.01
Cecum to Splenic Flexure	-0.61	(-0.84, -0.17)	0.01
Cecum to Rectum	-0.57	(-0.82, -0.11)	0.01





Highest Performing (Greater than Mean+0.5*SD)
Above Average (Mean to Mean+0.5*SD)
Below Average (Mean-0.5*SD to Mean)
Very Low Performing (Less than Mean-0.5*SD)



Highest Performing (Greater than Mean+0.5*SD)
Above Average (Mean to Mean+0.5*SD)
Below Average (Mean-0.5*SD to Mean)
Very Low Performing (Less than Mean-0.5*SD)

Editor's Notes

Background: Adenoma detection rate (ADR) and serrated polyp detection rate (SDR) vary significantly among colonoscopists. Colonoscopy inspection quality is the quality with which a colonoscopist inspects the colon for polyps and may explain some of this variation.

Findings: Overall colonoscopy inspection quality and individual colonoscopy inspection quality components (fold examination, luminal distension, and cleaning) correlate with ADR and SDR. Colonoscopists with similar ADRs and SDRs differ in their performance of the three colonoscopy inspection quality components.

Implications for Patient Care: Measuring colonoscopy inspection quality using video-recordings may facilitate identification of low performing colonoscopists in centers that cannot easily calculate ADR and SDR and could provide individual colonoscopists with targeted feedback to drive quality improvement efforts.