Adenoma detection rate is necessary but insufficient for distinguishing high versus low endoscopist performance CME

Hank S. Wang, MD, MSHS,[1,2,4,5,7] Joseph Pisegna, MD,[1,2,5] Rusha Modi, MD,[6] Li-Jung Liang, PhD,[3] Mary Atia, MD,[6] Minh Nguyen, MD,[6] Hartley Cohen, MD,[1] Gordon Ohning, MD, PhD,[1,2,5] Martijn van Oijen, PhD,[2,7] Brennan M.R. Spiegel, MD, MSHS[1,2,4,5,7]

Los Angeles, California, USA

Background: Endoscopist quality is benchmarked by the adenoma detection rate (ADR)—the proportion of cases with 1 or more adenomas removed. However, the ADR rewards the same credit for 1 versus more than 1 adenoma.

Objective: We evaluated whether 2 endoscopist groups could have a similar ADR but detect significantly different total adenomas.

Design: We retrospectively measured the ADR and multiple measures of total adenoma yield, including a metric called ADR-Plus, the mean number of incremental adenomas after the first. We plotted ADR versus ADR-Plus to create 4 adenoma detection patterns: (1) optimal (↑ ADR/ ↑ ADR-Plus); (2) one and done (↑ ADR/ ↓ ADR-Plus); (3) all or none (↓ ADR/ ↑ ADR-Plus); (4) none and done (↓ ADR/ ↓ ADR-Plus).

Setting: Tertiary-care teaching hospital and 3 nonteaching facilities servicing the same patient pool.

Patients: A total of 3318 VA patients who underwent screening between 2005 and 2009.

Main Outcome Measurements: ADR, mean total adenomas detected, advanced adenomas detected, ADR-Plus.

Results: The ADR was 28.8% and 25.7% in the teaching (n = 1218) and nonteaching groups (n = 2100), respectively (P = .052). Although ADRs were relatively similar, the teaching site achieved 23.5%, 28.7%, and 29.5% higher mean total adenomas, advanced adenomas, and ADR-Plus versus nonteaching sites (P < .001). By coupling ADR with ADR-Plus, we identified more teaching endoscopists as optimal (57.1% vs 8.3%; P = .02), and more nonteaching endoscopists in the none and done category (42% vs 0%; P = .047).

Limitations: External generalizability, nonrandomized study.

Conclusion: We found minimal ADR differences between the 2 endoscopist groups, but substantial differences in total adenomas; the ADR missed this difference. Coupling the ADR with other total adenoma metrics (eg, ADR-Plus) provides a more comprehensive assessment of adenoma clearance; implementing both would better distinguish high- from low-performing endoscopists. (Gastrointest Endosc 2013;77:71-8.)

Colonoscopy is the only screening test that is both cancer detecting and cancer preventing; its dual purposes are to detect prevalent colorectal cancer (CRC) and to clear the colon of suspected CRC precursors, ie, adenomatous polyps.[1] Because endoscopists cannot definitively predict which polyp is an adenoma,[2] much less which adenomas

Abbreviations: ADR, adenoma detection rate; WLAVA, West Los Angeles Veterans Administration; CRC, colorectal cancer.

DISCLOSURE: *All authors disclosed no financial relationships relevant to this publication. This research was supported by a Veterans Affairs Merit Award (IIR 08-310) to Dr Spiegel and an NIH training grant (T32DK07180-34) to Dr Wang.*

See CME section; p. 108.

0016-5107/$36.00
http://dx.doi.org/10.1016/j.gie.2012.08.038

Received April 19, 2012. Accepted August 30, 2012.

Current affiliations: Department of Medicine (1), Division of Gastroenterology, VA Greater Los Angeles Healthcare System, Department of Medicine (2), Division of Digestive Diseases, Division of General Internal Medicine and Health Services Research (3), David Geffen School of Medicine at UCLA, Department of Health Services (4), UCLA School of Public Health, Los Angeles, California, CURE Digestive Diseases Research Center (5), Division of General Internal Medicine (6), Cedars-Sinai Medical Center, UCLA/VA Center for Outcomes Research and Education (CORE) (7), Los Angeles, California, USA.

Presented at the ASGE Presidential Plenary at Digestive Disease Week 2011, May 19-22, 2012, Chicago, Illinois.

The opinions and assertions contained herein are the sole views of the authors and are not to be construed as official or as reflecting the views of the Department of Veteran Affairs.

Reprint requests: Brennan M.R. Spiegel, MD, MSHS, VA Greater Los Angeles Healthcare System, David Geffen School of Medicine at UCLA, UCLA/VA Center for Outcomes Research and Education (CORE), 11301 Wilshire Blvd., Bldg. 115, Room 215, Los Angeles, CA 90073.

will become malignant,[3,4] guidelines encourage endoscopists to remove every suspected adenoma, no matter its perceived risk for malignant transformation.[5-7]

The U.S. Multi-Society Colorectal Cancer Task Force currently supports the adenoma detection rate (ADR) as the quality indicator with which to benchmark endoscopists' ability to clear the colon of adenomas.[8] Defined as the proportion of a provider's screening colonoscopies in which 1 or more adenomas are removed, ADR has several advantages as a performance measure: it is simple to calculate, easy to understand, well studied,[9-11] and familiar to most endoscopists. Moreover, the ADR is currently the only indicator known to predict interval CRC.[12]

Although the ADR provides insight into endoscopists' ability to detect adenomas, it is limited as a quality indicator because it cannot distinguish between those who only find 1 versus more than 1 adenoma per case. An endoscopist who consistently finds 1 adenoma receives the same ADR credit as another who consistently finds more than 1 adenoma. Because every adenoma carries some risk of malignancy,[4] it is likely that these 2 endoscopists provide different levels of CRC protection, even though they receive equal ADR credit.

In contrast, it is possible for an endoscopist to have a low ADR yet find many adenomas. This endoscopist may find many adenomas at a time, but not find any adenomas at other times (ie, all or none adenoma detection). Following this logic, it is possible that endoscopists might fall into 1 of 4 performance categories. Table 1 demonstrates these 4 hypothetical patterns arranged in a 2 × 2 table. In the right upper quadrant are excellent endoscopists who maintain both a high ADR and high *total* adenoma detection. In the right lower quadrant are endoscopists who maintain a high ADR, yet do not achieve high total adenoma detection (ie, one and done). In the left upper quadrant are endoscopists who have a low ADR, yet find many adenomas. These all or none endoscopists do not find adenomas frequently, but when they find 1 adenoma, they may find more thereafter (high multiplicity rate). Finally, endoscopists in the left lower quadrant neither achieve a high ADR nor find many adenomas (ie, none and done); these are low-performing endoscopists.

If these patterns truly exist, then it would suggest that the ADR is necessary but insufficient to distinguish high-performing (right upper quadrant) from low-performing (left lower quadrant) endoscopists. Instead, it would be necessary to combine information from both the ADR and another metric of total adenoma detection to properly distinguish top from bottom performers.

In this study, we present data showing how 2 groups of endoscopists can have a similar ADR yet vary significantly in terms of total adenoma detection. We further demonstrate how the ADR is necessary but insufficient to characterize adenoma detection performance among endoscopists. We introduce a metric that captures adenoma detection beyond the first one found—the ADR-Plus—and compare its performance with other established metrics including mean total adenomas detected, advanced ADR, and multiplicity detection rate. Finally, we reveal how individual endoscopists could be stratified by ADR and ADR-Plus to categorize the 4 hypothesized adenoma detection patterns.

Take-home Message

- The adenoma detection rate (ADR), although well studied and easy to calculate, fails to distinguish endoscopists who find a high versus low total number of adenomas.
- Coupling ADR with a second measure of total adenoma detected (eg, ADR-Plus) yields 4 patterns of adenoma detection and provides a more comprehensive assessment of adenoma clearance.

METHODS

Study setting and patients

To evaluate and categorize adenoma detection patterns among a group of endoscopists, we retrospectively reviewed consecutive screening colonoscopies performed between January 2005 and June 2009 at the West Los Angeles Veterans Administration (WLAVA) Medical Center and a group of 3 nonteaching fee-for-service facilities contracted by WLAVA to assist with CRC screening. The WLAVA endoscopy unit is part of the University of California at Los Angeles Gastroenterology Training Program. Eighty percent of procedures are performed by a gastroenterology trainee with faculty oversight, and the remaining procedures are independently performed by faculty. During this study, there were 39 trainees and 12 faculty members who performed colonoscopies in the WLAVA unit.

The contracted nonteaching facilities are located 3, 20, and 100 miles from the WLAVA and service the same population of veterans awaiting screening colonoscopy. These nonhospital-based endoscopy units range in size from 2 to 7 gastroenterologists per site. Physicians and nurses at all sites were blinded to the purpose of the study. The WLAVA Institutional Review Board approved this study (PCC: 2010-010031).

Study procedures

Patients scheduled for screening colonoscopy received the option to choose between a teaching and nonteaching facility; there was no systematic process to centrally allocate patients. Because it is possible that patients might vary between sites, we adjusted for potential differences (see Statistical Analyses). Regardless of site, all patients attended a colonoscopy education class held at the VA. These classes provided uniform instruction regarding bowel preparation and proper preprocedural diet.

TABLE 1. Hypothetical endoscopist adenoma detection patterns

Low ADR				High ADR			
High total adenomas							
All or none				**Optimal**			
Case	**1st adenoma**	**2nd adenoma**	**3rd adenoma**	**Case**	**1st adenoma**	**2nd adenoma**	**3rd adenoma**
1				1			
2	X	X		2			
3				3	X	X	
4				4			
5				5			
6				6	X	X	X
7				7			
8	X	X	X	8			
9				9	X		
10				10			
ADR = 20%				ADR = 30%			
Total adenomas = 5				Total adenomas = 6			
Low total adenomas							
None and done				**One and done**			
Case	**1st Adenoma**	**2nd Adenoma**	**3rd Adenoma**	**Case**	**1st Adenoma**	**2nd Adenoma**	**3rd Adenoma**
1				1			
2				2	X		
3	X			3			
4				4			
5				5	X		
6				6			
7				7			
8				8	X		
9				9			
10				10			
ADR = 10%				ADR = 30%			
Total adenomas = 1				Total adenomas = 3			

The table demonstrates how endoscopists can be categorized by the ADR on 1 axis and total adenomas detected on another. In this manner, endoscopists may fall into 4 potential patterns: (1) high ADR and high total adenomas (optimal, right upper quadrant); (2) high ADR and low total adenomas (one and done, right lower quadrant); (3) low ADR and high total adenomas (all or none, left upper quadrant); and (4) low ADR and low total adenomas (none and done, left lower quadrant).
ADR, Adenoma detection rate.

In preparation for colonoscopy, patients received sodium phosphate (Fleet Phospho-soda; C.B. Fleet, Lynchburg, Va), magnesium citrate, or an oral lavage of polyethylene glycol. Patients were instructed to take the regimen the night before colonoscopy. All patients received their bowel preparation with directions from the centralized VA pharmacy. Subjects received sedation with midazolam and either fentanyl or meperidine or with a propofol drip, at the discretion of the individual site. The WLAVA Pathology Department diagnosed all tissue specimens for both cohorts; pathologists were blinded to whether specimens originated from teaching or nonteaching sites.

Adenoma detection metrics

We calculated a range of adenoma detection metrics on both the group level and for individual endoscopists who performed at least 30 procedures in the study, as follows:

Adenoma detection rate. We calculated the ADR by dividing the total number of screening procedures in which 1 or more histologically confirmed adenomas were detected by the total number of screening procedures performed.

Advanced ADR. We calculated the advanced ADR by dividing the total number of screening procedures in which advanced adenoma(s) were detected by the total number of screening procedures performed. We defined a case of an advanced adenoma as any with 3 or more adenomas of any size, 1 or more large adenomas (>1 cm), or 1 or more adenomas with villous architecture or high-grade dysplasia.[6]

Multiplicity detection rate. We calculated the multiplicity detection rate by dividing the total number of screening procedures in which 2 or more adenomas were detected by the total number of screening procedures performed. The multiplicity detection rate provides equal credit to those who find 2 adenomas per procedure versus more than 2 adenomas per procedure. This metric builds on the ADR but remains dependent on the ADR. That is, an endoscopist must find 1 adenoma before finding 2 or more.

Mean total adenomas detected. We calculated the mean total adenomas detected by dividing the total number of adenomas detected by the total number of screening procedures performed. As with the multiplicity detection rate, mean total adenomas detected incorporates and is dependent on the ADR.

ADR-Plus. We derived the ADR-Plus by calculating the mean number of adenomas found after the first in procedures in which 1 or more adenomas were detected. In this manner, ADR-plus is a true measure of incremental gain after the first adenoma detected and is independent of the ADR itself (unlike the multiplicity detection rate or mean total adenomas). For example, an ADR-Plus of 0.8 indicates that an endoscopist finds, on average, 80 additional adenomas after the first detected per 100 screening procedures in which at least 1 adenoma is found. In contrast to ADR-Plus, the ADR describes the proportion of procedures in which at least 1 adenoma was found, but it does not capture how many additional adenomas are detected. It is possible to maintain a high ADR yet not find any additional adenomas after the first one per procedure (ie, one and done pattern); this suboptimal pattern would remain undetected by the ADR. Although ADR-Plus is independent of the ADR, the mean includes all adenomas detected, including the first, and is not independent of the ADR.

Statistical analysis

We calculated descriptive statistics for demographic characteristics, patient risk factors, and colonoscopy outcomes (Table 1). We used 2-sample Student *t* tests and Wilcoxon rank sum tests for continuous parametric and nonparametric variables, respectively, and χ^2 tests for categorical variables between provider groups.

We compared the ADR, the advanced ADR, and the multiplicity detection rate between provider groups by using a χ^2 test. We then performed multivariable logistic regression to measure the impact of provider group on each metric while adjusting for patient age, type of sedation, type of preparation received, preparation quality, farthest point reached during colonoscopy (including cecal intubation), years of endoscopist experience, and presence of a trainee.

Moving beyond the ADR, we next compared the total adenoma yield among cohorts by using negative binomial multivariable regression, a technique suited for nonparametric count data such as adenoma detection, where the mean and variance are unequal.[13] We expressed the difference in adenoma yield between groups as an adjusted ratio of means (derived from e^{β}) with 95% confidence intervals. This analysis provided a summary estimate of the difference in overall adenoma detection between groups.

We then compared ADR-Plus between groups by using the Wilcoxon rank sum test. We report the relative difference in ADR-Plus between cohorts and compare the relative difference with that achieved by using other adenoma detection indicators, including the ADR, the advanced ADR, multiplicity detection rate, and mean total adenomas detected.

Finally, we categorized each participating endoscopist into 1 of 4 hypothesized behavioral categories. We profiled each endoscopist by using their individual ADR plotted against their ADR-Plus. This yielded a graph with 4 quadrants: (1) high ADR and high ADR-Plus (optimal performance); (2) high ADR and low ADR-Plus (one and done); (3) low ADR and high ADR-Plus (all or none); and (4) low ADR and low ADR-Plus (none and done). We used SAS System for Windows Version 9.2 (SAS Institute, Cary, NC) and STATA version 8.0 (StataCorp, College Station, Tex) for all analyses.

RESULTS

Patient characteristics and descriptive statistics

There were 1218 and 2100 patients in the teaching and nonteaching groups, respectively. There were differences between groups, as shown in Table 2, including variations in the use of phospho-soda, achievement of cecal intubation, and use of propofol. There was a nonsignificant trend toward differences in bowel preparation quality between groups.

TABLE 2. Screening colonoscopy summary data for teaching and nonteaching patients

Parameter	Teaching (N = 1218)	Nonteaching (N = 2100)	*P* value
Age, y, mean (SD)	63.9 (7.9)	62.9 (7.6)	<.001
Male sex, no. (%)	1181 (97.0)	2054 (97.8)	.13
Prep type, no. (%)			<.0001
MoviPrep	336 (28.7)	412 (19.7)	
Magnesium citrate	452 (38.6)	1527 (73.0)	
Fleet Phospho-soda	381 (32.6)	152 (7.3)	
Polyethylene glycol	1 (0.1)	1 (0.05)	
Type of sedation, no. (%)			<.0001
Propofol	2 (0.16)	1308 (62.3)	
Fentanyl and midazolam	149 (12.2)	8 (0.4)	
Meperidine and midazolam	928 (76.2)	589 (28.1)	
Other	139 (11.4)	195 (9.3)	
Cecal intubation rate, no. (%)			<.0001
All cases	1111 (91.2)	2056 (97.9)	
Aborted cases	67 (5.5)	17 (0.8)	
Prep quality, no. (%)			.08
Not poor	1062 (87.2)	1785 (85.0)	
Poor	156 (12.8)	315 (15.0)	
Endoscopist experience, y, mean (SD)	20.1 (9.5)	22.1 (13.6)	<.0001

Adenoma detection between groups

The total number of adenomas detected was significantly higher in the teaching group (mean 0.56 per case) than the nonteaching group (mean 0.43 per case) (P = .014). In negative binomial regression adjusting for differences between groups, the teaching site detected 44% more adenomas overall compared with the nonteaching sites (95% CI for relative difference, 1.3-82%; P = .002). The difference in total adenomas is further illustrated in Figure 1, which demonstrates the adenoma detection curves by provider group. Inspection of the overlapping curves reveals a relatively similar percentage of procedures with only 1 adenoma detected, but larger differences in the percentage of procedures with 2 to 8 adenomas detected. After the first adenoma was detected, there was a sharper dropoff in adenoma detection in the nonteaching versus teaching sites.

In unadjusted analysis, the overall ADR in the teaching and nonteaching groups was 28.8% and 25.7%, respectively (P = .052). After adjustment, patients in the teaching group had 37% higher odds of adenoma detection (odds ratio 1.37; 95% CI, 1.1-1.7; P = .006). There were also more advanced adenoma procedures in the teaching (13.2%) versus nonteaching (9.4%) groups (P = .0007). After adjustment, patients in the teaching group had 53% higher odds of advanced adenoma detection (odds ratio 1.53; 95% CI, 1.13-2.08; P = .006). In unadjusted analysis, the multiplicity detection rates in the teaching and nonteaching groups were 13.5% and 9.4%, respectively (P = .0003); this difference remained significant in regression analysis. Finally, the ADR-Plus in the teaching and nonteaching groups was 0.93 and 0.66 incremental adenomas after 1, respectively (P = .0002).

Comparing measures of incremental adenomas detected after the first

To quantify adenoma yield after the first one detected, we compared metrics between groups including mean total adenomas detected, advanced ADR, multiplicity detection rate, and ADR-Plus (Table 3). Although the relative difference in the ADR was only 10.8% between teaching and nonteaching providers, the relative differences in mean, multiplicity detection rate, advanced ADR, and ADR-Plus were 23.5%, 30%, 28.7%, and 29.5%, respectively, indicating that these alternative metrics outperformed ADR in distinguishing between groups.

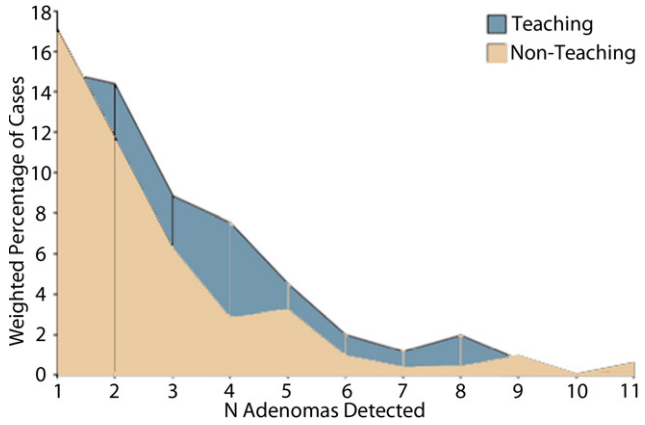

Figure 1. Adenoma detection curves for teaching and nonteaching groups. Inspection of the overlapping curves reveals a similar percentage of procedures with only 1 adenoma detected, but larger differences in the percentage of procedures with 2 to 8 adenomas detected. After the first adenoma was detected, there was a sharper dropoff in adenoma detection in the nonteaching than the teaching sites. When calculating integrated shaded areas under these weighted curves (eg, finding 2 polyps is weighted twice finding 1 polyp), the area is exactly equal to the mean total adenomas detected or 0.557 adenomas versus 0.427 adenomas for teaching versus nonteaching groups, respectively ($P = .014$). In this manner, the curves provide a useful heuristic for demonstrating the mean adenoma results between groups.

Categorizing endoscopist by ADR versus ADR-Plus

We calculated ADR and ADR-Plus among 19 endoscopists (7 teaching and 12 nonteaching) who completed at least 30 colonoscopies during the study period. When applying the U.S. Multi-Society CRC Task Force quality benchmark of a 25% or higher ADR (for male patients), only 50% of nonteaching endoscopists met this criterion; in contrast, 86% of teaching endoscopists met the Task Force ADR quality indicator ($P = .012$). We further stratified endoscopists by ADR-Plus, a measure of incremental adenoma detection independent of the ADR. Using the mean ADR-Plus among all endoscopists as a threshold (0.8 incremental adenomas) and crossing with the U.S. Multi-Society Task force ADR threshold of 25%, we created a graph with 4 adenoma detection quadrants (Fig. 2). Inspection of Figure 2 demonstrates a wide distribution among endoscopists within the optimal, one and done, all or none, and none and done categories. The figure visually demonstrates how endoscopists meeting the Task Force ADR criterion can still vary widely in terms of total adenomas detected, here measured with ADR-Plus. For example, we found more teaching than nonteaching endoscopists in the optimal quadrant (57.1% vs 8.3%, $P = .02$) despite both groups having relatively similar ADRs.

DISCUSSION

The ADR is the currently endorsed quality indicator for adenoma detection; the metric is meant to accurately reflect endoscopists' ability to clear the colon of adenomas during screening colonoscopy.[14] However, the ADR is an imperfect metric because it provides no information about incremental adenomas detected after the first. In this study, we reveal how an endoscopist can maintain an acceptable ADR but find far fewer adenomas than other endoscopists with the same ADR. When further stratifying endoscopists by a metric that accounts for incremental adenomas beyond the first detected (by using ADR-Plus), we demonstrate how endoscopists can distribute across 4 potential adenoma detection patterns (Fig. 2).

Because every adenoma carries some risk for malignant transformation and because guidelines currently support the removal of every polyp detected,[5-7] our data suggest that measuring the ADR alone does not fully capture endoscopist performance, that is, the ADR appears necessary, but is not sufficient. Instead, our results suggest that the ADR should be coupled with a second measure of total adenomas detected. Although the relative difference in the ADR between groups was just more than 10%, the relative differences in mean total adenomas, multiplicity detection rate, advanced ADR, and ADR-Plus were 23.5%, 30%, 28.7%, and 29.5%, respectively.

In Figure 2, we plot the ADR versus ADR-Plus for each endoscopist in the study who performed at least 30 or more procedures. By the ADR alone, 86% and 50% of teaching and nonteaching endoscopists, respectively, were categorized as performing optimally, assuming the U.S. Multi-Society CRC Task Force ADR threshold of 25% or higher ($P = .12$). By coupling the ADR with ADR-Plus, however, we further stratified these endoscopists to identify not only those who find a high proportion of procedures in which 1 or more adenomas are found (high ADR), but also to identify those who find a greater number of incremental adenomas after the first detected (high ADR-Plus). We found more teaching than nonteaching endoscopists in the optimal quadrant despite both groups having relatively similar ADRs.

We observed that the difference in adenoma detection between groups was largely driven by the difference in finding 2, 3, and 4 adenomas per procedure (Fig. 1). Compared with the teaching group, the nonteaching group exhibited a greater decrease in the percentage of procedures across this range. The diverging curves indicate that adenoma detection varied significantly after removal of the first adenoma; the ADR missed this difference.

This study has limitations. First, it is limited to a cohort of overwhelmingly male U.S. veterans; generalizability beyond this population remains untested. Nonetheless, this cohort provided a natural experiment well suited for proof-of-concept testing of ADR and alternative metrics. Whether the groups are teaching versus nonteaching, veterans versus nonveterans, or men versus women, we

TABLE 3. Difference in adenoma detection between teaching and nonteaching groups

	Teaching (n = 1218)	Nonteaching (n = 2100)	Relative difference between groups	*P* value
ADR (proportion with ≥1 adenoma detected)	28.8	25.7	10.8	.052
Multiplicity detection rate (proportion with ≥2 adenomas detected)	13.5	9.4	30.0	.0003
Mean total adenomas detected per procedure	0.56	0.43	23.5	.014
Advanced ADR (proportion with advanced adenomas detected)	13.2	9.4	28.7	.0007
ADR-Plus (mean total adenomas detected beyond first)	0.93	0.66	29.5	.0002

Although the difference in the ADR between groups was relatively small (10.8% relative difference), the difference in total adenomas detected was significantly different between groups, regardless of the metric used. The ADR-Plus, multiplicity detection rate, and advanced ADR maximally distinguished between groups, with 30% relative difference in values between groups.
ADR, Adenoma detection rate.

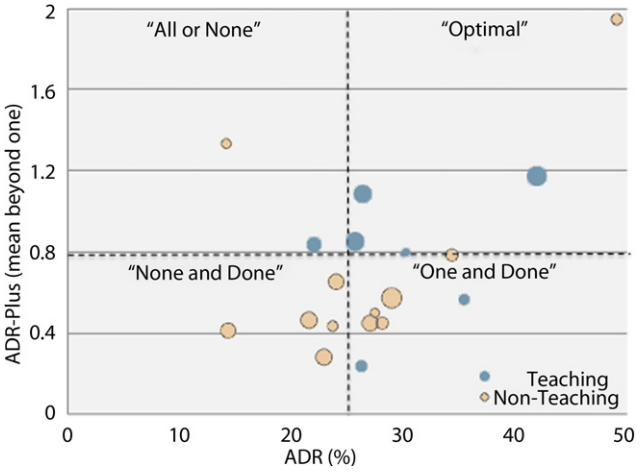

Figure 2. Distribution of endoscopists among 4 adenoma detection patterns. The figure demonstrates a wide distribution among endoscopists within the optimal, one and done, none and done, and all or none categories. The figure reveals how endoscopists meeting the U.S. Multi-Society CRC Task Force ADR Criterion (>25% for male patients) can still vary widely in terms of total adenomas detected, here measured with ADR-Plus, a metric independent of the adenoma detection rate (ADR). There were more teaching than nonteaching endoscopists in the optimal quadrant (57.1% vs 8.3%, $P = .02$) despite both groups having relatively similar ADRs. The size of the circle reflects the total number of procedures performed by the endoscopist.

found 2 groups of endoscopists with nearly the same ADR, but a notably different total number of adenomas detected. In so doing, we detected variations between groups, and these variations might be caused by shortcomings in the ADR itself, the very reason for considering a second metric of total adenomas detected, such as mean adenomas or ADR-Plus. This difference exposes problems with the ADR that total adenoma metrics such as ADR-Plus may remedy, independent of the patient population. Although we acknowledge that the practicality of capturing all these measures, including ADR-Plus, is not trivial.

Another limitation is that we found differences in patient and provider characteristics between groups. However, we performed adjusted analyses to account for these differences and found that adjustment only amplified, not contracted, the observed differences in adenoma detection. Moreover, although differences between groups might explain ADR variations, they would not account for the even larger relative difference in the total adenoma metrics.

The ADR remains an important performance measure because it is the only quality indicator linked to incident CRC.[12] This distinguishes it from any other measuring, including the ADR-Plus, because it is supported by longitudinal data that the metric truly predicts interval cancer. Nonetheless, the ADR appears to have important limitations. In this study, we reveal how 2 provider groups with similar ADRs can have very different total adenoma detection. We conclude that the ADR is necessary to accurately profile endoscopist performance, but that the ADR alone is not sufficient to fully measure screening colonoscopy quality. Because current guidelines support the removal of every adenoma, no matter its risk for malignant transformation, we believe that the combination of ADR with a total adenoma metric is superior to measuring the ADR alone. Future research should evaluate whether using supplemental measures, such as ADR-Plus, better predicts interval cancer rates than the ADR alone.

REFERENCES

1. Lieberman D. Screening/early detection model for colorectal cancer. Why screen? Cancer 1994;74:2023-7.
2. Lawrance IC, Sherrington C, Murray K. Poor correlation between clinical impression, the small colonic polyp and their neoplastic risk. J Gastroenterol Hepatol 2006;21:563-8.
3. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology 1990;98:371-9.
4. Risio M. The natural history of adenomas. Best Pract Res Clin Gastroenterol 2010;24:271-80.
5. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence. Gastroenterology 2003;124:544-60.

6. Rex D, Johnson D, Anderson J, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. Am J Gastroenterol 2009;104:739-50.
7. Davila RE, Rajan E, Baron TH, et al. ASGE guideline: colorectal cancer screening and surveillance. Gastrointest Endosc 2006;63:546-57.
8. Rex DK, Bond JH, Winawer S, et al. Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the U.S. Multi-Society Task Force on Colorectal Cancer. Am J Gastroenterol 2002;97:1296-308.
9. Chan MY, Cohen H, Spiegel BM. Fewer polyps detected by colonoscopy as the day progresses at a Veteran's Administration teaching hospital. Clin Gastroenterol Hepatol 2009;7:1217-23.
10. Barclay RL, Vicari JJ, Doughty AS, et al. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006;355:2533-41.
11. Rogart JN, Siddiqui UD, Jamidar PA, et al. Fellow involvement may increase adenoma detection rates during colonoscopy. Am J Gastroenterol 2008;103:2841-6.
12. 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.
13. Cameron A, Trivedi P. Regression analysis of count data. Cambridge: Cambridge University Press; 1998.
14. Williams JE, Faigel DO. Colonoscopy reports and current state of performance measures. Gastrointest Endosc Clin N Am 2010;20:685-97.