Thieme

Which is the better polyp detection metric: adenomas per colonoscopy or adenoma detection rate? A simulation modeling study

Authors
Todd MacKenzie[1,2], Sikai Xiao[3], William H Hisey[4,5], Christina M Robinson[4,5], Lynn Butterly[4], Joseph C Anderson[6,7,8]

Institutions
1 Department of Biomedical Data Science, The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, United States
2 Department of Medicine, Dartmouth College Geisel School of Medicine, Hanover, United States
3 The Analytics Institute, Dartmouth-Hitchcock Medical Center, Lebanon, United States
4 Department of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, United States
5 New Hampshire Colonoscopy Registry, Lebanon, United States
6 Dartmouth College Geisel School of Medicine, Hanover, United States
7 White River Junction VA Medical Center, White River Junction, United States
8 University of Connecticut School of Medicine, Farmington, United States

Key words
Endoscopy Lower GI Tract, Polyps / adenomas / ..., CRC screening, Statistics

received 16.7.2024
accepted after revision 18.9.2024
accepted manuscript online 14.10.2024

Bibliography
Endosc Int Open 2024; 12: E1366–E1373
DOI 10.1055/a-2417-6248
ISSN 2364-3722

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

Corresponding author
Dr. Joseph C Anderson, Dartmouth College Geisel School of Medicine, Hanover, United States
Joseph.Anderson@dartmouth.edu

Supplementary Material is available at https://doi.org/10.1055/a-2417-6248

ABSTRACT

Background and study aims We compared the ability of adenoma detection rate (ADR) and adenoma per colonoscopy (APC) to assess endoscopist detection, using statistical principles and simulations.

Patients and methods We simulated a population of endoscopists and patients to compare the ability of ADR versus APC for capturing true endoscopist ability (TEA). We compared these rates with and without adjustment for patient and exam factors using multivariable models, and adjustment for imprecision due to low volume using empirical Bayes (shrinkage). Power calculations were used to compare the ability of ADR and APC to distinguish higher from lower rates over two time periods for an endoscopist.

Results APC and ADR had similar discriminatory ability for assessing TEA. This increased with higher volumes and after adjusting for risk factors and low volume using shrinkage. Higher APC and ADRs had higher power for comparing endoscopist detection over two time periods, but APC was superior to ADR. For example, there was 29% power to distinguish APCs (n = 200 colonoscopies) 0.10 from 0.15, similar to the power (28%) to distinguish corresponding ADRs: 10% and 14%. However, at same volume (n = 200), the power to distinguish higher APC rates (0.50 vs.0.75) was greater (89%) than the power (78%) for corresponding ADRs (39% vs.53%).

Conclusions Adjusting for patient and exam factors and/or using shrinkage techniques for lower-volume endoscopists can increase the correlation between TEA for both ADR and APC. For higher detection rates, APC offers more power than ADR in distinguishing differences in detection ability.

Introduction

Adenoma detection rate (ADR) is a validated colonoscopy quality measure [1, 2, 3, 4] that measures the proportion of screening colonoscopies with an adenoma [5–7.] ADR is a surrogate for the adenoma miss rate, with low ADRs suggesting higher miss rates [5, 6, 7, 8]. Patients who have endoscopists with high ADR are at lower risk for post-colonoscopy colorectal cancer (PCCRC) than patients whose endoscopists have low ADR [1, 2, 9]. However, ADR does not account for presence of multiple adenomas.

Another quality indicator, adenomas per colonoscopy (APC) [10, 11, 12, 13, 14, 15, 16, 17], represents the average number of adenomas detected per screening colonoscopy. APC credits endoscopists who clear the colon of all precursor lesions beyond the "one and done" reflected in the ADR [7, 13, 18, 19]. There is more information captured in APC than ADR. ADR is analogous to the proportion of games in which a basketball player scores; APC is analogous to points per game. However, the average points per game per professional basketball player is close to 10 whereas the average number of adenomas detected per exam is less than one, making the latter statistic more difficult to grasp intuitively. In particular, the median reported APC for published data from two large population-based registries have been less than one [17, 20].

The relative ability of ADR and APC to measure endoscopist quality is not directly estimable without looking at long-term outcomes such as CRC. This would be important when comparing detection over two time periods to determine whether an endoscopist is improving. A reasonably powered study examining CRC requires over a decade and thousands of patients. Thus, alternative methods for comparing the power of ADR versus APC for differentiating the detection ability of two endoscopists is important, especially in clinical practice.

Detection rates do not account for case-mix variation among endoscopists, such as patient age, gender, and exam indication. Some endoscopists may also have low volumes. Experts have proposed including diagnostic and surveillance exams in order to simplify calculations for busy endoscopists as well as to increase power by increasing the number of colonoscopies in the calculation [21]. Data from the New Hampshire Colonoscopy Registry (NHCR) have demonstrated that there could be differences based on exam indication, screening versus surveillance, as well as patient demographics such as smoking or sex [22, 23, 24].

In this study, we used a statistical framework including simulations to compare how well ADR and APC measure true endoscopist ability (TEA). First, we used power calculations to compare two hypothetical datasets of either two endoscopists with different abilities or a single endoscopist across two time periods using ADR and APC. Second, we compared the degree to which ADR and APC captured the actual ability of different endoscopists to detect adenomas using a simulation of a population of endoscopists and patients. In addition, we evaluated these rates after accounting for case-mix adjustment and small-volume sampling variation using the statistical approach of shrinkage (empirical Bayes).

Methods

This study was not a statistical analysis of any particular dataset. It was an evaluation of statistical methods commonly used to evaluate endoscopist ability using statistical and mathematical principles including simulations.

Population and sample used as basis for simulations

Our evaluation of the statistical properties of APC in comparison to ADR are based on data from the NHCR. Our previous papers have provided details about the NHCR database [10, 11, 23, 25, 26, 27]. Briefly, individuals who have a colonoscopy in the state of New Hampshire voluntarily consent to participate in the NHCR (IRB: Committee for the Protection of Human Subjects, CPHS# 15834). Patients also complete an NHCR Patient Questionnaire prior to colonoscopy, including demographic, health behavior, and personal and family history data. Endoscopists and/or endoscopy nurses complete the NHCR Colonoscopy Procedure Form during or immediately after colonoscopy. Data on this form include indication of exam, and the location, size, and method of resection or biopsy for all findings [27].

Our simulations and analyses are based on NHCR patients with known adequate bowel preparation, complete exams to the cecum, no personal history of inflammatory bowel disease or CRC and with complete patient demographic and risk information. ▶ **Fig. 1** demonstrates the flow of each analysis.

Adenoma detection and adenoma per colonoscopy calculations

Evaluation 1: Comparing APC to ADR using simulations

Simulation of endoscopists and their detection ability

The first step in our simulation addresses endoscopist detection ability. We will refer to it as TEA and define it for each endoscopist as the probability of detecting an adenoma if it exists in the colon of a patient they are examining. As shown in **Supplementary Table 1**, the TEAs in our simulation were assumptions that varied between 35% and 99% according to a uniform distribution (each point equally likely) among endoscopists. They did not depend on any endoscopist characteristics such as years of experience, because that was not part of our simulation.

Although improvements in technique [28] or changes in bowel prep procedures can improve detection rates [25, 29], we make the simplifying assumption that TEA is constant over time the same between within and between patients. In other words, we assume that an endoscopist's probability of detecting adenomas is constant and does not change for each adenoma if there is more than one adenoma in a patient or change from one patient to the next, although the endoscopist may find more in some patients because they have more adenomas in their colon, which we address in Step B.

We simulated two scenarios. In the first, representing calculation of APC and ADR from 1 year of an endoscopy group, the number of endoscopists was 82. In the second scenario, the number of endoscopists was 136, representing an endoscopy group from several years. These counts match the number of

endoscopists in the NHCR who performed colonoscopies on 30 or more patients in the most recent year and during all the years of the registry, respectively. **Supplementary Table 1** shows quantiles of exam volumes for each endoscopist (range of 30 to 360) for 1 year of results as well as for all years (range of 30 to 5490).

Simulation of colonoscopies with varying adenoma prevalence in patients and detection ability in endoscopists

In this second step, we simulated patients and their colons with varying numbers of adenomas, depending on patient characteristics, sex, age, as well as three types of exams: screening, surveillance and diagnostic. Patient sex and age and their variation between endoscopist panels were simulated according to the distributions reflecting the NHCR described in **Supplementary Table 1** (i.e., patient age and sex vary both within and between endoscopists). We varied the frequency of screening, surveillance and diagnostic exams between endoscopists using the distributions listed in **Supplementary Table 1** based on the NHCR. The proportion of screening, surveillance, and diagnostic exams averaged 56%, 25%, and 19% respectively, and varied between the endoscopists with standard deviations of 12%, 14% and 15%, respectively.

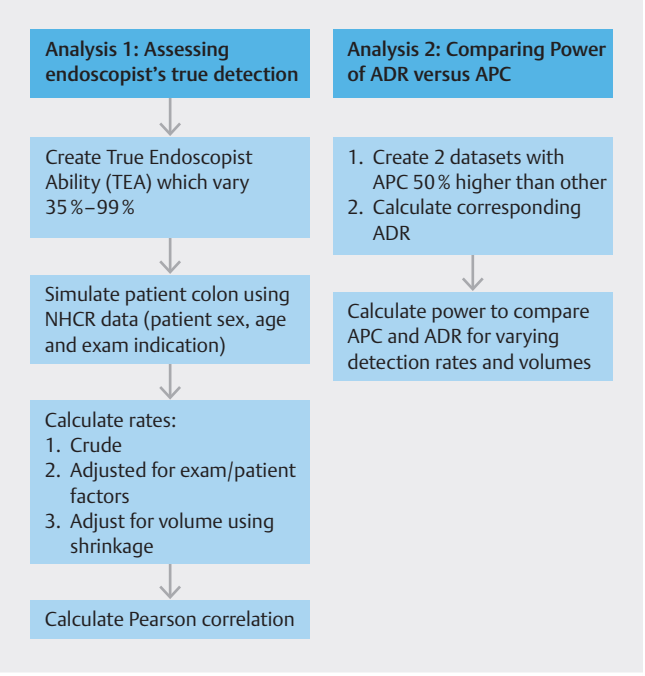
To create the colons of simulated patients, we randomly generated the number of adenomas in each patient's colon using a gamma-Poisson (i.e. negative binomial) model specified in **Supplementary Table 1**, which reflects the relationship between the number of adenomas and exam indication, patient age, and patient sex. We set the intercept of the adenomas-per-patient model to align the overall average of adenomas in the simulated population with the ADR and APC found in the NHCR, after accounting for probability of detecting each adenoma. The adenomas detected in each colonoscopy were simulated using a binomial distribution, Bin(N,p) with N being the number of actual adenomas in the patient, and p equal to the probability of detection by the endoscopist (TEA).

Calculation of detection rates

The third step was to calculate detection rates using the simulated endoscopist and patient data. We calculated both the ADR and APC three ways: 1) conventional (unadjusted); 2) adjusted for patient sex, patient age, and exam indication; and 3) adjusted for imprecision due to low volumes using shrinkage (empirical Bayes). The purpose of shrinkage is to reduce variation due to small samples (volumes), allowing for more reliable rate calculations for endoscopists with low exam volume [30]. The adjustment is accomplished using a multivariable logistic regression model followed by calculation of the observed over expected, followed by a rescaling to the overall rates. The adjusted and shrunken version is calculated using a multivariable logistic regression model with a random intercept for endoscopist. The shrunken version is based on the mean posterior prediction for each endoscopist. Note: We did not implement and evaluate shrinkage in the simulations in which each endoscopist has the same volume because in that case, shrinkage cannot change endoscopist rankings.

Correlation of ADR and APC with TEA

In the fourth and final step, we evaluated the ability of ADR and APC to discriminate low from high performing endoscopists by calculating the Pearson correlation. The Pearson correlation decreases with increasing sampling error (e.g. lower endoscopist volume) and with the amount of unadjusted confounding.



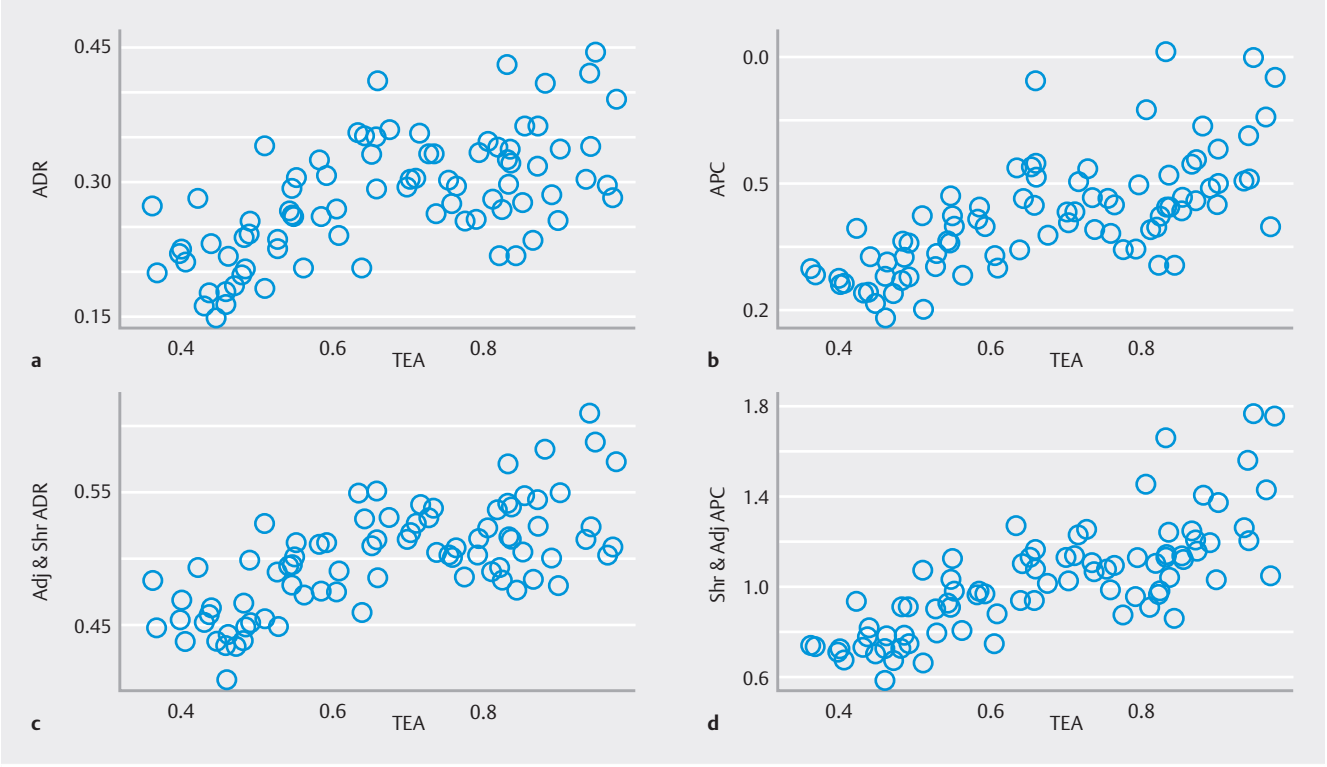
▶ **Fig. 1** Flow chart for each analysis.

Simulation code

The simulations were conducted using the statistical software R. The code is available in the appendix. This allows interested readers to utilize the code and modify the simulation settings to reflect characteristics of a different panel of endoscopists and patients.

Evaluation 2: Power comparison of APC and ADRdr for difference between two endoscopists

We calculated the power to compare various values of APC between two time periods for an endoscopist or to compare two endoscopists. While endoscopist-level APC is usually calculated by dividing the number of adenomas detected for each endoscopist by the total number of exams for that endoscopist, we created hypothetical APCs. Specifically, we made the assumption that the ratio between the two detection rates was fixed at 1.5 (for example, APC 1 = 0.4 and APC 2 = 0.6) and that each endoscopist's count of detected adenomas among all patients follows a Poisson distribution. We calculated power for different levels of APC and volumes of colonoscopies. ADRs for this exercise were not based on clinical data but, instead, were calculated based on an assumed Poisson distribution. Specifically, although ADR for each endoscopist is usually calculated as the number of colonoscopies with at least one adenoma divided by the total number of colonoscopies, respective ADRs in this analysis were calculated based on the statistical principle that the count of detected adenomas follows a Poisson distribution



▶ **Fig. 2** Scatterplots. **a** Crude ADR vs true endoscopist ability (TEA). **b** Crude APC vs TEA under a single iteration (of 500 in total) based on 1 year of data (median endoscopist volume of 90) and **c** and **d** all years of data (median endoscopist volume of 600).

with mean of λ, then the probability of detecting at least one adenoma is 1-exp(-λ).

Statistically, our power calculation to compare the APCs between endoscopist is based on use of a two-sample *t*-test, which is asymptotically equivalent to use of the score test from a Poisson regression. The power to compare the ADR uses the same underlying Poisson distributed counts that have been dichotomized into "one or more adenomas" versus "no adenomas." The comparison of the ADR between two endoscopists uses a Pearson chi-square test. The power is calculated using the proportion in which the test statistic yields a $P < 0.05$.

Results

Evaluation 1: Simulations to assess discriminatory ability of APC and ADR for assessing individual endoscopist detection ability

▶ **Fig. 2** shows the scatterplots resulting from one iteration of the simulations based on a typical single year in the NHCR. ▶ **Fig. 2a** shows the conventional ADR versus the endoscopist's true ability to detect adenomas (TEA). Each point represents one endoscopist. ▶ **Fig. 2b** shows APC versus TEA. ▶ **Fig. 2c** and ▶ **Fig. 2d** show the corresponding plots for ADR and APC after they are adjusted to account for varying case mix between endoscopists and shrunken to account for small sample (low volume) variation. Essentially, the simulation results reported below arise from generating these plots 1000 times and averaging the correlations.

▶ **Table 1** presents the strength of the association of TEA with ADR and APC from our simulation, using either 1 year of data, or all years of data, from an endoscopy patient panel similar to the NHCR as shown in **Supplementary Table 1**. In addition to ADR and APC (i. e. conventional or crude rates), we report findings for adjusted and adjusted empirical Bayes (shrinkage) versions of the ADR and APC. The Pearson correlation of TEA with crude APC based on 1 year of results, 0.77, is marginally better than the correlation of TEA with crude ADR of 0.73. There is little difference in the Pearson correlations of TEA with the adjusted empirical Bayes, (0.89 for APC and 0.88 for ADR), based on 1 year of data. However, there is an increase in Pearson correlations for each metric, APC and ADR, when using shrunken method along with adjusted.

▶ **Table 1** also shows simulations using all years of NHCR data in which the median endoscopist volume is almost 600 exams. This analysis also shows a small advantage of APC in comparison with ADR. The Pearson correlation of TEA with the crude APC and crude ADR are 0.85 and 0.83, respectively. The Pearson correlation of TEA with the adjusted APC and ADR are 0.91 and 0.87, respectively.

The final simulation results presented in ▶ **Table 1** reflect a simulation in which 100 endoscopists each conduct the same number of endoscopies, either 50, 200, or 1000. As above, APC performs incrementally better than ADR, whether it is the crude or adjusted version for each volume level.

▶ **Table 1** Correlation of APC and ADR with ability to discriminate underlying endoscopist ability.

		Pearson correlation with TEA	
Time frame	**Method**	**APC**	**ADR**
One year in NHCR - 82 endoscopists with volumes from 30 to 436	Conventional	0.77	0.73
	Adjusted	0.81	0.77
	Adjusted and shrunken	0.89	0.88
All years of NHCR – 136 endoscopists with volumes from 30 to 5486	Conventional	0.85	0.83
	Adjusted	0.91	0.87
	Adjusted and shrunken	0.92	0.90
Same volume in 100 endoscopists*			
50 endoscopies each	Conventional	0.66	0.61
	Adjusted	0.69	0.63
200 endoscopies each	Conventional	0.83	0.80
	Adjusted	0.88	0.84
1000 endoscopies each	Conventional	0.90	0.89
	Adjusted	0.97	0.95

Crude, adjusted and shrunken APC and ADR based on either 1 year or 5 years of a simulated
colonoscopy database similar in size to the New Hampshire Colonoscopy Registry.
*Shrinkage not performed here because it will not change rates differently among endoscopists
because all physicians have the same volume.
TEA, true endoscopist ability.

Analysis 1: Power comparison of APC and ADR for difference between two endoscopists

▶ **Table 2** displays the power for detecting a 50% difference in detection ability between two endoscopists or time periods for a single endoscopist using APC and ADR. Using the actual count of adenomas from each colonoscopy (APC) as opposed to dichotomizing the count as at least one versus none (ADR) results in a modest increase in power to discriminate between two endoscopists. For instance, if the underlying ADRs of two endoscopists are 30% and 45% or if an endoscopist has an increase in ADR of 30% to 45% over two time periods, the power to detect a difference between these two rates based on 200 colonoscopies by each endoscopist or for each time period is 69% using the APC and 61% using the ADR. ▶ **Table 2** shows that power for both ADR and APC increases with higher volumes of colonoscopies. In addition, there is an increase in power for higher rates of APC and ADR.

Discussion

Higher ADRs have been shown to be protective of PCCRC [5, 6, 7, 8]. Because APC counts the number of all adenomas detected as opposed to their presence or absence as in ADR, it may be a better measure of an endoscopist's ability to detect all adenomas. While there are data which examine the association between both ADR and APC and PCCRC[1, 2, 7], there are no data which directly compare these two metrics with respect to assessing endoscopists' true detection abilities. In our analyses, we used mathematical and probabilistic principles to compare ADR and APC.

In our first set of analyses, we used statistical simulation to compare APC and ADR in capturing the TEA, defined as the probability that an endoscopist will identify an adenoma if it exists in the colon. Detection rates are surrogates for the more important measure, adenoma miss rate (AMR).[5, 6, 7, 8]. It may not be clear that the miss rate of an endoscopist is addressed in this simulation. In fact, the true AMR of an endoscopist is one minus what we define as TEA. That is, the probability that an endoscopist misses an adenoma is 1 – TEA. If more than one adenoma existed, we assumed for the sake of simplicity that each detection had the same probability. In these simulations, we observed that APC had a slight advantage over ADR. At most volumes, the APC outperformed the ADR by a difference of less than 0.05 in its Pearson correlation with TEA when calculated using a conventional approach (crude or simple). The same was true when we compared the shrunken and case mix-adjusted APC with the shrunken and case mix-adjusted ADR.

A relevant secondary observation is that both the APC and the ADR were more effective at measuring TEA when they were adjusting for case mix. This case-mix adjustment may be important to quality improvement in endoscopy. We are not aware of any endoscopy groups that calculate detection rates adjusted for case mix, aside from sex [5, 6, 7, 8]. Potential other factors to consider include age, body mass index (BMI), and smoking, which are known significant neoplasia risk factors [23, 24].

We also observed that shrinkage statistical methods (empirical Bayes) also improved the correlation with TEA of both APC and ADR, especially when the endoscopist had fewer exams, such as in the calculation of detection rates for the most recent year. One of the authors of this paper has used this approach to compare outcomes for coronary artery bypass graft surgery among centers with different volumes of surgeries [30]. Specifically, they observed that using random effects for centers yields shrinkage estimates reduces the likelihood of falsely classifying a small-volume center as having exceptional or very poor results. We think that this approach deserves further attention from endoscopy groups that report endoscopist-level detection rates.

▶ **Table 2** Power to detect a difference in detection ability of two datasets*.

APC		ADR		Number of endoscopies	Power using APC	Power using ADR
Dataset A	Dataset B	Dataset A	Dataset B			
0.10	0.15	10%	14%	50	11%	10%
0.20	0.30	18%	26%	50	17%	15%
0.30	0.45	26%	36%	50	23%	20%
0.40	0.60	33%	45%	50	29%	24%
0.50	0.75	39%	53%	50	35%	27%
0.60	0.90	45%	59%	50	41%	30%
0.10	0.15	10%	14%	200	29%	28%
0.20	0.30	18%	26%	200	52%	47%
0.30	0.45	26%	36%	200	69%	61%
0.40	0.60	33%	45%	200	81%	71%
0.50	0.75	39%	53%	200	89%	78%
0.60	0.90	45%	59%	200	93%	82%
0.10	0.15	10%	14%	500	61%	58%
0.20	0.30	18%	26%	500	89%	85%
0.30	0.45	26%	36%	500	97%	94%
0.40	0.60	33%	45%	500	99%	98%
0.50	0.75	39%	53%	500	100%	99%
0.60	0.90	45%	59%	500	100%	100%
0.10	0.15	10%	14%	1000	89%	87%
0.20	0.30	18%	26%	1000	99%	99%
0.30	0.45	26%	36%	1000	100%	100%
0.40	0.60	33%	45%	1000	100%	100%
0.50	0.75	39%	53%	1000	100%	100%

*Either two time periods for an endoscopist or two endoscopists) using the APC versus using the ADR, across varying APC (0.10–0.90) and ADR (10%-59%) levels and numbers of colonoscopies (50–1000).
ADR, adenoma detection rate; APC, adenoma per colonoscopy.

In our second set of analyses, we used statistical power to compare the ability of ADR versus APC to detect differences in colonoscopy quality between two simulated data sets with different volumes. These data sets could be either two time periods for a single endoscopist or for two endoscopists. In this analysis, one dataset had an APC which was 50% higher than the other data set. Not surprisingly, higher volumes were associated with higher power for both ADR and APC to detect differences in adenoma detection. At these high rates, both ADR and APC had 100% power.

These data have implications for practicing endoscopists. Published data suggest that a minimum volume of 500 colonoscopies might be needed to calculate an ADR with appropriately narrow 95% confidence intervals [31]. We observed that at a volume of 500 colonoscopies or greater, the power for ADR and APC was high for distinguishing superior endoscopists with higher detection rates. For example, a volume of 500 exams for each time period would provide a power of 94% for differentiating two ADRs of 26% and 36%. An ADR of 25% is the current benchmark for practicing endoscopists [8]. If an endoscopist improved from 18% to 26% in order to meet the current benchmark, a volume of 500 colonoscopies would provide 85% to determine if there was a true increase in ADR for this endoscopist.

We recently published data demonstrating that an APC of 0.5 or higher in endoscopists with an adequate ADR of 25% or higher was associated with a lower hazard for post colonoscopy as compared with those with an APC < 0.5 and an adequate ADR [17]. If an endoscopist improves from an APC of 0.4 to 0.6 over two time periods, each time period would need more than 200 exams in each in order to provide > 90% power to determine if there was an improvement in adenoma detection. At this volume, the power for APC (81%) and ADR (71%) are both lower than 90%.

Another observation was that for fixed exam volumes, APC provided superior power to detect differences between two endoscopists or two time periods for one endoscopist, one with a 50% higher rate of detection, as compared with the corresponding ADRs. However, the difference was the highest for higher detection rates. For example, when comparing two endoscopists with 200 colonoscopies, the difference between the power for APC and ADR was 11% (APC 93% and ADR 82%) for endoscopists with APC of 0.6 and 0.9 and corresponding ADR of 45% and 59%. As highlighted above for a high volume of 1000 exams and/or high rates of ADR and APC, power was 100% for both measures. Conversely for lower APC (0.10, 0.15) and corresponding ADR of 10% and 14%, the difference was only 1% (APC 29% and ADR 28%). This is because ADR is very similar to APC if the APC is low. For instance, if the APC is 0.20, the corresponding ADR would be 0.18, as shown in ▶ **Table 1**.

We acknowledge that there are some clinical concerns regarding use of APC in practice, including the potential increased burden for endoscopists with respect to counting all adenomas that are removed in each patient. APC use might incentivize physicians to place individual polyps in separate bottles, which could increase cost. One solution might be to photo-document all lesions (with artificial intelligence assistance) to allow determination of the actual number of adenomas [32]. Another limitation would be the potential burden of collecting data for the statistical approach for case-mix adjustment. One solution could be use of natural language processing, which is being used in many hospitals to collect patient data.

We acknowledge that there is no perfect way to determine an endoscopist's true ability. For instance, even tandem colonoscopy is unable to fully estimate true ability because both endoscopists in a tandem colonoscopy may miss an adenoma [33]. We also recognize that our simulation settings did not account for changes in an endoscopist's detection ability, which may not be reflective of clinical practice where physicians' detection rates may be improving due to technology or technique [25, 28]. However, because there have been no studies using statistical and mathematical principles to compare the ability of APC versus ADR to assess endoscopists' ability to detect adenomas, our results provide novel insight into this issue. Our simulation included exams that were complete and performed on patients who had adequate bowel preparation. We based our simulation on patient sex and age as well as exam indication. However, we did not include BMI and sedation, which might influence adenoma prevalence. Other strengths include use of a statewide population-based colonoscopy registry to inform our assumptions for patients and endoscopists.

Conclusions

Our analyses suggest that APC improvement in discrimination between two endoscopists is associated with their detection ability. Our findings do not necessarily require action. However, for those groups who only report ADR, our findings suggest that switching to reporting APC might be preferable but with only a slight advantage. For lower levels of ADR, APC may add little to the ability to statistically distinguish the detection ability of two endoscopists. However, it is important to note that the power can differ based on volume as well detection rates. For example, when examining a volume of 200 colonoscopies (10 per week for 6 months), APC has more power than ADR (69% versus 61%) for clinically relevant ADR changes of 26% to 36%. Thus, at these rates, APC may have an advantage at colonoscopy volumes that are clinically relevant.

Overall, however, for higher levels of ADR, APC may offer more power to differentiate between endoscopists or between two time periods for one endoscopist. When detection rates are very high, especially in high exam volume scenarios, both ADR and APC have 100% power for discrimination. We also that observed APC may not have a substantially higher correlation with individual TEA than ADR. However, adjusting for case-mix factors such as patient characteristics and exam indication and/or using shrinkage techniques for lower-volume endoscopists can increase the correlation between TEA and both ADR and APC.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding Information

Division of Cancer Prevention, National Cancer Institute
http://dx.doi.org/10.13039/100007316
5R01CA243449
Optimizing colorectal cancer prevention: a multi-disciplinary, population-based investigation of serrated polyps using risk prediction and modeling; Grant Recipient: Lynn F. Butterly, M.D.

References

[1] Corley DA, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2024; 370: 1298–1306

[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: 1795–1803 doi:10.1056/NEJMoa0907667

[3] Kaminski MF, Robertson DJ, Senore C et al. Optimizing the quality of colorectal cancer screening worldwide. Gastroenterology 2020; 158: 404–417 doi:10.1053/j.gastro.2019.11.026

[4] Anderson JC, Butterly LF. Colonoscopy: quality indicators. Clin Transl Gastroenterol 2015; 6: e77 doi:10.1016/j.gie.2024.04.2905

[5] 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 doi:10.1111/j.1572-0241.2002.05812.x

[6] Rex DK, Petrini JL, Baron TH et al. Quality indicators for colonoscopy. Gastrointest Endosc 2006; 63: S16–S28 doi:10.1016/j.gie.2006.02.021

[7] Rex DK. Detection measures for colonoscopy: considerations on the adenoma detection rate, recommended detection thresholds, withdrawal times, and potential updates to measures. J Clin Gastroenterol 2020; 54: 130–135 doi:10.1097/MCG.0000000000001301

[8] Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. Gastrointest Endosc 2015; 81: 31–53

[9] Cross AJ, Robbins EC, Saunders BP et al. Higher adenoma detection rates at screening associated with lower long-term colorectal cancer incidence and mortality. Clin Gastroenterol Hepatol 2022; 20: e148–e167

[10] Anderson JC, Butterly LF, Weiss JE et al. Providing data for serrated polyp detection rate benchmarks: an analysis of the New Hampshire Colonoscopy Registry. Gastrointest Endosc 2017; 85: 1188–1194 doi:10.1016/j.gie.2017.01.020

[11] Anderson JC, Hisey W, Mackenzie TA et al. Clinically significant serrated polyp detection rates and risk for postcolonoscopy colorectal-cancer data from the New Hampshire Colonoscopy Registry. Gastrointest Endosc 2022; 96: 310–317

[12] Denis B, Sauleau EA, Gendre I et al. The mean number of adenomas per procedure should become the gold standard to measure the neoplasia yield of colonoscopy: a population-based cohort study. Dig Liver Dis 2014; 46: 176–181

[13] Gessl I, Waldmann E, Penz D et al. Evaluation of adenomas per colonoscopy and adenomas per positive participant as new quality parameters in screening colonoscopy. Gastrointest Endosc 2019; 89: 496–502

[14] Hilsden RJ, Bridges R, Dube C et al. Defining benchmarks for adenoma detection rate and adenomas per colonoscopy in patients undergoing colonoscopy due to a positive fecal immunochemical test. Am J Gastroenterol 2016; 111: 1743–1749 doi:10.1038/ajg.2016.449

[15] Kumar AR. Set higher adenomas per colonoscopy benchmark. Gastrointest Endosc 2014; 80: 539–541 doi:10.1016/j.gie.2014.04.003

[16] Wang S, Kim AS, Church TR et al. Adenomas per colonoscopy and adenoma per positive participant as quality indicators for screening colonoscopy. Endosc Int Open 2020; 8: E1560–E1565 doi:10.1055/a-1261-9074

[17] Anderson JC, Rex DK, Mackenzie TA et al. Endoscopist adenoma per colonoscopy detection rates and risk for post colonoscopy colorectal cancer: data from New Hampshire Colonoscopy Registry. Gastrointest Endosc 2024; 99: 787–795

[18] Fedewa SA, Anderson JC, Robinson CM et al. Prevalence of 'one and done' in adenoma detection rates: results from the New Hampshire Colonoscopy Registry. Endosc Int Open 2019; 7: E1344–E1354

[19] Vennelaganti S, Cuatrecasas M, Vennalaganti P et al. Interobserver agreement among pathologists in the differentiation of sessile serrated from hyperplastic polyps. Gastroenterology 2021; 160: 452–454 e1

[20] Wieszczy Pm Bugajski M, Januszewicz W et al. Comparison of quality measures for detection of neoplasia at screening colonoscopy. Clin Gastroenterol Hepatol 2023; 21: 200–209 e6 doi:10.1016/j.cgh.2022.03.023

[21] Ladabaum U, Shepard J, Mannalithara A. Adenoma and serrated lesion detection by colonoscopy indication: The ADR-ESS (ADR Extended to all Screening/Surveillance) Score. Clin Gastroenterol Hepatol 2021; 19: 1873–1882 doi:10.1016/j.cgh.2021.04.027

[22] Anderson JC, Butterly LF, Goodrich M et al. Differences in detection rates of adenomas and serrated polyps in screening versus surveillance colonoscopies, based on the new hampshire colonoscopy registry. Clin Gastroenterol Hepatol 2023; 11: 1308–1312

[23] Anderson JC, Weiss JE, Robinson CM et al. Adenoma detection rates for screening colonoscopies in smokers and obese adults: Data from the New Hampshire Colonoscopy Registry. J Clin Gastroenterol 2017; 51: e95–e100

[24] Anderson JC, Calderwood AH, Christensen BC et al. Smoking and other risk factors in individuals with synchronous conventional high-risk adenomas and clinically significant serrated polyps. Am J Gastroenterol 2018; 113: 1828–1835

[25] Anderson JC, Butterly LF, Robinson CM et al. Impact of fair bowel preparation quality on adenoma and serrated polyp detection: data from the New Hampshire colonoscopy registry by using a standardized preparation-quality rating. Gastrointest Endosc 2014; 80: 463–70

[26] Anderson JC, Robinson CM, Butterly LF. Increased risk of metachronous large serrated polyps in individuals with 5- to 9-mm proximal hyperplastic polyps: data from the New Hampshire Colonoscopy Registry. Gastrointest Endosc 2020; 92: 387–393 doi:10.1016/j.gie.2020.04.034

[27] Greene MA, Butterly LF, Goodrich M et al. Matching colonoscopy and pathology data in population-based registries: development of a novel algorithm and the initial experience of the New Hampshire Colonoscopy Registry. Gastrointest Endosc 2011; 74: 334–340 doi:10.1016/j.gie.2011.03.1250

[28] Butterly L, Robinson CM, Anderson JC et al. Serrated and adenomatous polyp detection increases with longer withdrawal time: results from the New Hampshire Colonoscopy Registry. Am J Gastroenterol 2014; 109: 417–426

[29] Butterly LF, Nadel MR, Anderson JC et al. Impact of colonoscopy bowel preparation quality on follow-up interval recommendations for average-risk patients with normal screening colonoscopies: Data from the New Hampshire Colonoscopy Registry. J Clin Gastroenterol 2020; 54: 356–364

[30] MacKenzie TA, Grunkemeier GL, Grunwald GK et al. A primer on using shrinkage to compare in-hospital mortality between centers. Ann Thorac Surg 2015; 99: 757–761

[31] Do A, Weinberg J, Kakkar A et al. Reliability of adenoma detection rate is based on procedural volume. Gastrointest Endosc 2013; 77: 376–380

[32] Rex DK, Hardacker K, MacPhail M et al. Determining the adenoma detection rate and adenomas per colonoscopy by photography alone: proof-of-concept study. Endoscopy 2015; 47: 245–50 doi:10.1055/s-0034-1391330

[33] Rutter CM, Nascimento de Lima P, Lee JK et al. Too good to be true? Evaluation of colonoscopy sensitivity assumptions used in policy models Cancer Epidemiol Biomarkers Prev 2022; 31: 775–782