Adenoma Detection Rate in Asymptomatic Patients with Positive Fecal Immunochemical Tests

Eugene Kligman, MD [1], Wenfang Li [2], George J. Eckert, MAS [2], Charles Kahi, MD, MSc [1,3]

[1] Department of Medicine, Indiana University School of Medicine; [2] Department of Biostatistics, Indiana University School of Medicine and Richard M. Fairbanks School of Public Health, [3] Richard L. Roudebush VA Medical Center; Indianapolis, Indiana

Author contributions: E Kligman: acquisition of data; analysis and interpretation of data; drafting of the manuscript, Wenfang Li: analysis and interpretation of data; critical revision of the manuscript for important intellectual content, G Eckert: analysis and interpretation of data; critical revision of the manuscript for important intellectual content, C Kahi: study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content

No conflicts of interest

This study was presented as an oral abstract at Digestive Disease Week 2017 (Chicago, Illinois)

Corresponding author:
Charles J. Kahi, MD, MSc, FASGE
Professor of Clinical Medicine
Indiana University School of Medicine
Gastroenterology Section Chief, Roudebush VA Medical Center
1481 W 10th street, 111G
Indianapolis, IN 46202
Phone: (317) 988-3682
Fax: (317) 988-5313
E-mail: ckahi2@iu.edu

_____

ABSTRACT

Background and Aims: The adenoma detection rate (ADR) is a powerful measure of screening colonoscopy quality. Patients who undergo colonoscopy for evaluation of a positive fecal immunochemical test (FIT) have increased prevalence of colorectal neoplasia, but it is not known whether separate quality benchmarks are required. The aim of this study was to compare the conventional ADR to the ADR of colonoscopies performed for the evaluation of positive FIT, in asymptomatic average-risk patients.

Methods: Patients ≥ 50 years old who underwent colonoscopy for the evaluation of a positive FIT between 1/1/2013 and 7/31/2014 at a tertiary Veterans Affairs Medical Center were identified. FIT performed for any indication other than average-risk screening were excluded. The comparison group included average-risk patients ≥ 50 years old undergoing screening colonoscopy during the same time frame. The 2 groups were compared for ADR, advanced neoplasm (adenoma ≥ 10 mm, tubulovillous, high-grade dysplasia, CRC, sessile serrated polyp [SSP] ≥ 10 mm), CRC, and SSP detection after propensity score adjustment using a logistic regression model adjusted for endoscopist.

Results: There were 207 patients in the FIT group, and 601 in the screening colonoscopy comparison group. After propensity score adjustment, ADR (72.9% vs. 50.0%, p=0.003), number of adenomas per colonoscopy (3.3 ± 3.6 vs. 1.4 ± 2.3, p=0.033) and advanced neoplasm detection rate (32.4% vs. 11.0%, p<0.0001) were significantly higher in the FIT group. There were no significant differences in number of CRC, and the SSP detection rate.

Conclusions: In this cohort of average-risk Veterans, the ADR of colonoscopies performed for evaluation of a positive FIT was higher than the ADR of screening colonoscopies. Patients with a positive FIT also had significantly more adenomas per colonoscopy and advanced neoplasms. These findings suggest that the quality of colonoscopies performed for a positive FIT is insufficiently assessed by the conventional ADR, and requires additional quality metrics.

INTRODUCTION

A fundamental goal of effective colonoscopy is colorectal cancer (CRC) prevention through the detection and removal of precancerous polyps. There is considerable evidence showing that the effectiveness of colonoscopy depends on the quality of the examination. The adenoma detection rate (ADR), defined as the proportion of screening colonoscopies where at least one adenoma is found, is currently the prime surrogate measure of colonoscopy performance quality [1], and is inversely associated with the risk of post-colonoscopy CRC[2,3]. The American Society for Gastrointestinal Endoscopy (ASGE)/American College of Gastroenterology (ACG) Task Force on Quality in Endoscopy has set minimum target detection rates in average-risk individuals of 30% for men and 20% for women [1]. Although the ADR is a robust quality metric, unanswered questions remain about the optimal method for measurement. The conventional ADR calculation is based on screening colonoscopies, and minimum detection rates for colonoscopies performed for other indications, notably the evaluation of occult gastrointestinal bleeding, are not well defined.

In its guideline, the ASGE/ACG Task Force states that further research is needed to determine whether (and within which parameters) the ADR can be used as a quality metric in screening programs that use fecal occult blood or immunochemistry testing to select patients for colonoscopy[1]. This issue is important because fecal immunochemical tests (FIT) have now emerged as the non-invasive CRC screening test of choice, and are the frontline CRC screening modality in several countries[4]. Patients with positive FIT have high prevalence rates of adenomas, advanced adenomas, and cancer[5], adding to the rationale for development of separate ADR targets. In a recent guideline, the US Multi-Society Task Force on Colorectal Cancer proposed ADR benchmarks of 45% in men and 35% in women on colonoscopy examinations performed to evaluate a FIT-positive test (using a threshold of 20 μg hemoglobin/gm dry stool or less). However, this was a weak recommendation based on very low quality evidence, derived from series reporting the PPV for adenoma detection in FIT-based

3

screening programs[6]. In the Veterans Affairs healthcare system, colonoscopy and FIT are available concurrently as primary CRC screening options, allowing an opportunity to compare the yield of colonoscopy performed for both indications in the same population. Such data are a requisite step to develop ADR benchmarks for FIT-based programs. The aim of this study is to compare the conventional ADR to the ADR of colonoscopies performed for the evaluation of positive FIT, in asymptomatic average-risk Veterans.

METHODS

The study was approved by the Institutional Review Board at Indiana University-Purdue University at Indianapolis and by the Richard L. Roudebush VA Medical Center Research and Development Committee in Indianapolis, Indiana. Data sources included the Roudebush VA Medical Center GI endoscopy electronic database (ProVation Medical; Minneapolis, Minn), and the Veterans Affairs Vista Computerized Patient Record System (CPRS).

We conducted a retrospective study to compare the ADR for colonoscopies performed to evaluate a positive FIT, to the conventional ADR derived from screening colonoscopies. We included asymptomatic, average-risk Veterans ≥ 50 years old who had undergone outpatient colonoscopy between January 1, 2013 and July 31, 2014. In one group, the colonoscopy was done for CRC screening and in the other group, the colonoscopy was done for evaluation of a positive FIT, initially performed for CRC screening. We excluded patients with overt gastrointestinal bleeding, any indication other than screening (such as iron deficiency anemia, diarrhea, etc.), a personal history of CRC or inflammatory bowel disease, a significant family history of CRC (first degree relative ≤ 60 years old or multiple relatives with CRC), and those who had undergone colonoscopy for any indication in the preceding 5 years. The FIT in use at the Roudebush VAMC is the qualitative Polymedco (Cortland, NY) FIT, with cutoff set at 20 μg hemoglobin/gm dry stool. Of the total 435 patients who had a positive FIT during the study time frame, 207 met inclusion criteria. We randomly selected 601 average-risk patients

(approximating a 3:1 ratio) who had undergone screening colonoscopy during the same time frame, using the same eligibility criteria as for the positive FIT group.

Data collection and definitions

For each eligible procedure, we collected data regarding patient demographics and physical characteristics (including age, sex, race, weight, height), cigarette smoking, alcohol use, Charlson Comorbidity Score[7], medication use (including calcium, NSAIDs and statins), sedation type and dose, bowel preparation quality (based on the Aronchick scale), trainee (GI fellow) participation, cecal intubation, withdrawal time, and resected polyp characteristics (including location, size in millimeters, morphology, and histology). The attending physician for each procedure was also recorded. Cecal intubation was defined as explicit documentation in the report that the appendix, cecal landmarks, and ileocecal valve were visualized and photographed. Advanced neoplasms included advanced adenomas (≥1 cm, villous histology, and high-grade dysplasia), sessile serrated polyps (SSP) ≥ 1cm, and cancer.

Based on these polyp data, the following measurements were derived: ADR (proportion of colonoscopies with at least one adenoma, including advanced neoplasms), number of adenomas per colonoscopy (APC), advanced neoplasm detection rate (proportion of colonoscopies with at least one advanced neoplasm), and SSP detection rate (proportion of colonoscopies with at least one SSP).

Statistical Analysis

Patient and procedure characteristics were summarized for the FIT and screening colonoscopy groups. Counts (percentages) were presented for categorical variables and mean ± standard deviation (SD) for continuous variables. Categorical characteristics were compared with chi-square tests. A Mantel-Haenszel chi-square test was used to compare ADR between the 2 groups. Two-sample *t* tests were used to compare continuous variables.

Given the retrospective nature of the study design, and potential for selection bias and confounding, we used propensity score analysis. This method approximates randomized

controlled trial situations, to allow comparison of outcomes between intervention groups[8].

Propensity scores were calculated using a logistic regression model adjusted for each

endoscopist, to predict group membership from the set of characteristics. Variables were

calculated for each group and compared by outcome for each of 5 strata. The variables were

BMI, age, Withdrawal Time (min), Total Time (min), Time of Assessment, attending, gender,

race, Current Smoker, on Statin, on NSAID, on Calcium, Charlson Comorbidity Index, Fellow

(trainee), Prep Quality, and cecal intubation. Outcomes included ADR, adenoma per

colonoscopy, advanced neoplasm detection rate, CRC, and SSP detection rate.

Finally, we performed an exploratory analysis to calculate ADR benchmarks for FIT-

positive colonoscopies, using the methodology described by Hilsden et al [9]. This was done

using data from endoscopists with at least 5 colonoscopies in each group, and linear regression

and calculation of Pearson's correlation coefficient using ADR- FIT as the outcome and ADR-

screening colonoscopy as the predictor.

The study was sufficiently powered (>80%) to detect the ADR differences observed

between the two groups. However the number of endoscopists with at least 5 colonoscopies in

each group was small, so only large correlations between the ADRs in the FIT-positive and

screening colonoscopy groups would be detected as statistically significant.

RESULTS

A total of 207 patients were included in the FIT group, and 601 in the screening

colonoscopy comparison group. The procedures were performed by a total of 23 endoscopists.

There were no significant differences between the groups with regards to age, race, cigarette

smoking, bowel preparation quality, trainee participation, and NSAID use. There were significant

differences in sex, body mass index, Charlson comorbidity score, withdrawal time, and calcium

supplementation (Table 1).

After propensity score adjustment, the ADR of the FIT group was significantly higher than the ADR of the screening colonoscopy group (72.9% vs. 50.0%, p=0.003). The APC (3.3 ± 3.6 vs. 1.4 ± 2.3, p=0.033) and advanced neoplasm detection rate (32.4% vs. 11.0%, p<0.0001) were significantly higher in the FIT group. The number of CRC and SSP detection rate were not significantly different (Table 2).

The ADRs and APCs of the 7 endoscopists who had performed at least 5 colonoscopies in each group are presented in Table 3. The exploratory analysis to calculate ADR benchmarks for FIT-positive colonoscopies could not be completed due to weak correlation between ADRs in the FIT-positive and screening colonoscopy groups (r=0.16).

DISCUSSION

In this cohort of average-risk Veterans, the ADR of colonoscopies performed for evaluation of a positive FIT was significantly higher than the conventional screening colonoscopy ADR. In addition, patients with a positive FIT had significantly more adenomas per colonoscopy and advanced neoplasms. In the FIT-colonoscopy group, the overall ADR was about 1.5 times higher, the APC for the endoscopists with the highest procedure volume were 2-3 fold higher, and the proportion of colonoscopies with at least one advanced neoplasm was 3 times higher than in the screening colonoscopy group. The ADR increase in the FIT-positive group was notable, given the high ADR in the comparison screening colonoscopy group (50%). These findings justify separate quality benchmarks, and suggest that the ADR target of 45% for men proposed by the ACG/ASGE Quality Task Force may in fact be an underestimate. Perhaps more importantly, the ADR alone may be insufficient to completely evaluate the quality for FIT-colonoscopy, given the significant increment in total adenoma and advanced neoplasia detection.

Previous studies [10-12] have shown that the APC is less corruptible and is a more comprehensive measure of colonoscopy quality than the conventional ADR. The ADR may mask information about overall adenoma burden, and endoscopists with seemingly similar ADRs may vary significantly with regards to total adenoma detection. The APC overcomes the "one and done" phenomenon[10], where an endoscopist could meet benchmark ADR targets by finding and removing one adenoma, but conduct a low quality examination subsequently. Minimum APC targets for screening colonoscopy have been proposed[13], but have not yet been guideline-endorsed. There are several reasons for this: (1) the APC metric is more cumbersome to measure than the conventional ADR, which could represent a significant barrier for uptake, (2) APC measurement may require placing resected adenomas into separate specimen bottles, which could increase costs, and (3) there are no data associating the APC with the risk of post-colonoscopy CRC as there are for the conventional ADR[2,3]. This reasoning may need to be reconsidered in the case of colonoscopy for positive FIT. In our study, the APCs for FIT-colonoscopies (range 2.5 to 4.6) were 2-3 fold higher than for screening colonoscopies performed by the same group of endoscopists (range 0.8-2.0), an incremental yield which would have been overlooked by measurement of the ADR alone.

The advanced neoplasia detection rate has also been proposed as a complementary quality metric[14], because advanced adenomas have a higher risk of progressing to CRC, and because the high ADRs reported in the recent literature may be inflated by increased detection of less clinically relevant small and diminutive adenomas. However, measurement is more cumbersome and challenging than the conventional ADR, due to concerns about variability in polyp size estimation among endoscopists, and variability in interpretation of the presence of villous elements and grade of dysplasia among pathologists. Our study revealed a 3-fold increase in advanced neoplasia detection rates in FIT-colonoscopy compared to screening

colonoscopy, which also challenges the notion that the conventional ADR is sufficient to capture colonoscopy quality in FIT-based screening programs.

One of our study's strengths is the use of propensity score analyses. Patients referred for CRC screening using colonoscopy or FIT may differ in significant ways; for example, some clinicians may prefer the non-invasive FIT in older frail patients, and/or be more likely to recommend colonoscopy for younger and healthier patients. Propensity score analysis reduces the confounding and biases inherent to the retrospective design and lack of baseline randomization[8]. An important consideration is whether there is residual bias due to endoscopists conducting a more careful examination in FIT-positive patients, due to heightened suspicion of the presence of colorectal neoplasia. However, procedure and withdrawal time were included in the propensity score matching, and the withdrawal times for colonoscopies in which no polyps were found were not significantly different between the 2 groups.

Our study has several limitations. First, it was conducted at a single tertiary VA medical center, with a predominantly male population. Studies conducted in different settings and populations are required to assess the generalizability of our findings. Second, the overall sample size was relatively small, and limited by the number of positive screening FIT-colonoscopies which were performed in average-risk patients during the study time frame. This resulted in small numbers of procedures per endoscopist, which likely contributed to the weak correlation between ADRs in the FIT-positive and screening colonoscopy groups and precluded the determination of FIT-colonoscopy ADR benchmarks as described by Hilsden et al.[9]. Third, we could not compare the FIT-positive group to patients who had undergone colonoscopy and had known negative FIT; in other words, the FIT status of the screening colonoscopy patients was unknown. However, the main comparison had to be against primary screening colonoscopy (without prior FIT), because this is how the conventional ADR is defined and measured. In addition, at our Medical Center, the two primary CRC screening options are either FIT or

colonoscopy, and average-risk patients who have a negative FIT are not referred to screening colonoscopy. Hence, addressing this issue requires a prospective study specifically designed with a FIT-negative colonoscopy group, which was beyond the scope of our work.

Additional considerations pertain to the fact that the FIT used at our facility is qualitative and set at 20 μg/g; hence the findings may be most applicable to settings using the same parameters. It is known that the performance characteristics of FIT depend on the brand, whether it is quantitative of qualitative, and on the positivity threshold of the latter[15]. These considerations have been elegantly reviewed elsewhere[5]. In general, quantitative FIT may be better due to automation of the interpretation process and the ability to adjust the positivity threshold. In addition, a lower cutoff is associated with increased sensitivity for colorectal cancer (CRC). However, in the United States, the Food and Drug Administration requires the use of qualitative FIT, and the 20 μg/g threshold is commonly chosen to balance increased neoplasia detection with colonoscopy availability. With these considerations, the US Multi-Society Task Force in fact suggests a cutoff of ≤ 20 μg/g. Thus our approach, based on a widely used qualitative FIT with a lower (but pragmatic) cutoff, is in agreement with FDA guidance and guideline recommendations, while being applicable and relevant to many practices which use FIT for average-risk CRC screening.

Another consideration is that adenoma yield would be expected to be increased for FIT-positive colonoscopies, raising the question whether the increased ADR reflects the quality of the FIT, rather than colonoscopy. In reality, the effect of a FIT-based screening strategy depends on several factors, including the performance characteristics of the FIT itself, the quality of the underlying program including administration of the test, completion rates, processing, appropriate colonoscopic follow-up, and critically, the quality of the colonoscopy which is performed to evaluate the patient with a positive test. The US Multi-Society Task Force[5] has in fact suggested quality metrics for each of these measures, in particular

colonoscopy quality for which ADR targets of 45% in men and 35% in women were suggested (if the FIT cutoff is ≤ 20 µg/g). In our study, all of the aforementioned factors are controlled for, given that the FIT used is qualitative and set at the appropriate threshold, and its use as part of a well-established screening program with high-quality processes for test administration and follow-up. The remaining unaddressed issue was whether the ADR and other quality metrics differed significantly for FIT-positive colonoscopies, compared to the conventional screening ADR; this was the primary aim of our work.

In conclusion, in a group of average-risk Veterans, the ADR of colonoscopies performed for evaluation of a positive FIT was higher than the ADR of screening colonoscopies. Patients with a positive FIT also had significantly more adenomas per colonoscopy and advanced neoplasms. Our findings suggest that the quality of colonoscopies performed for a positive FIT is insufficiently assessed by the conventional ADR, and requires concomitant measurement of the total number of adenomas per colonoscopy and advanced neoplasia detection rates. Additional studies in different settings are required to validate and determine the generalizability of our findings.

REFERENCES

1.   Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. Gastrointest Endosc 2015;81:31-53.
2.   Corley DA, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;370:2541.
3.   Kaminski MF, Regula J, Kraszewska E, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med;362:1795-803.
4.   Rex DK, Boland CR, Dominitz JA, et al. Colorectal Cancer Screening: Recommendations for Physicians and Patients From the U.S. Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2017;153:307-23.

5.   Robertson DJ, Lee JK, Boland CR, et al. Recommendations on fecal immunochemical testing to screen for colorectal neoplasia: a consensus statement by the US Multi-Society Task Force on colorectal cancer. Gastrointest Endosc 2017;85:2-21 e3.

6.   Jensen CD, Corley DA, Quinn VP, et al. Fecal Immunochemical Test Program Performance Over 4 Rounds of Annual Screening: A Retrospective Cohort Study. Ann Intern Med 2016;164:456-63.

7.   Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases 1987;40:373-83.

8.   D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in medicine 1998;17:2265-81.

9.   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;111:1743-9.

10.   Wang HS, Pisegna J, Modi R, et al. Adenoma detection rate is necessary but insufficient for distinguishing high versus low endoscopist performance. Gastrointest Endosc 2013;77:71-8.

11.   Lee TJ, Rutter MD, Blanks RG, et al. Colonoscopy quality measures: experience from the NHS Bowel Cancer Screening Programme. Gut 2012;61:1050-7.

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;46:176-81.

13.   Kahi CJ, Vemulapalli KC, Johnson CS, Rex DK. Improving measurement of the adenoma detection rate and adenoma per colonoscopy quality metric: the Indiana University experience. Gastrointest Endosc 2014;79:448-54.

14.   Greenspan M, Rajan KB, Baig A, Beck T, Mobarhan S, Melson J. Advanced adenoma detection rate is independent of nonadvanced adenoma detection rate. Am J Gastroenterol 2013;108:1286-92.

15.   Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med 2014;160:171.