Calculating the adenoma detection rate in screening colonoscopies only: Is it necessary? Can it be gamed?

Authors
Douglas K. Rex, Prasanna L. Ponugoti

Institution
Indiana University School of Medicine, Division of Gastroenterology/Hepatology, Indianapolis, USA

accepted after revision 21.5.2017

Bibliography

Corresponding author
Douglas K. Rex, MD, 550 North University Blvd, Suite 4159, Indianapolis, IN 46202, USA
Fax: +1-317-944-5449
drex@iu.edu

ABSTRACT
Background Current recommendations are to calculate the adenoma detection rate (ADR) in screening colonoscopies only. The need to confine the measure to screening has not been established.

Introduction
The adenoma detection rate (ADR) is the percent of patients age 50 or older undergoing screening colonoscopy who have one or more conventional adenomas detected and removed [1, 2]. However, the original proposal and definition of ADR did not specify that the calculation should be made in screening examinations only [3]. Rather, the original proposal for ADR measurement was for colonoscopies in persons age 50 and older, without reference to screening or diagnostic or surveillance examinations [3]. In 2006, the target for measuring ADR was narrowed to screening examinations by the joint quality task force of the American College of Gastroenterology/American Society for Gastrointestinal Endoscopy (ACG/ASGE) [1]. The more narrow definition was a logical step for the task force to take, because the original targets for an acceptable ADR of 25% in men and 15% in women were derived from screening colonoscopy studies that were available in 2002 [4–7]. Therefore, the task force deemed it most appropriate to narrow the measure to screening examinations only.

However, the need to narrow the ADR to screening examinations only has never been established. Narrowing the denominator makes the calculation of ADR more tedious, reduces the number of examinations available for assessment, and prolongs the period needed to establish ADR with a narrow confidence limit.

In addition, it seems that confining ADR to screening examinations creates the opportunity to corrupt ADR during prospective collection of ADR measurements. Most quality measures for colonoscopy, including detection measures, are subject to some degree of gaming. Common sense, in combination with a detailed understanding of the colonoscopy and colon polyp literature, suggests that some measures of polyp detection are more subject to gaming than others. For example, the well-known difficulties in accurately measuring colon polyp size and the absence of a widely available measuring tool for polyp

Methods
We retrospectively assessed our quality database for whether calculating ADR from screening, surveillance, and diagnostic colonoscopies (overall ADR) would alter conclusions about the performance of colonoscopists, compared to using an ADR based only on screening colonoscopies. We also prospectively tested the extent to which one physician could corrupt the screening-only ADR by changing the procedure indication after reviewing the examination findings.

Results
For 15 physicians, screening ADRs differed from the overall ADR by a mean of 2.6 percentage points (range 0–6.9 percentage points). Using the overall ADR rather than screening ADR changed the ADR from just below to just above the recommended screening threshold for one physician. In the prospective assessment, a single expert colonoscopist utilized indication gaming in patients with both screening and diagnostic indications and was able to increase his apparent screening-only ADR from 48.4% to 55.1%.

Conclusions
Use of an overall ADR rather than screening-only ADR could simplify ADR measurement, increase the number of examinations available to measure ADR, seldom affect whether a doctor meets recommended ADR thresholds, and eliminate the potential for gaming the ADR by changing the colonoscopy indication.
size, suggests that prospective application of a detection measure for “advanced adenomas” (AADR) would be severely affected by gaming [3]. Colonoscopists would be incentivized to upsize polyps to the ≥10 mm size category. Similarly, expert panels have resisted adoption of the “polyp detection rate” (PDR) because common sense suggests that in prospective applications it would incentivize the removal of clinically insignificant lesions, e.g. distal colon diminutive hyperplastic polyps [3].

ADR is also subject to gaming, though arguably less than either AADR or PDR. The principal mechanism for gaming ADR about which authors have raised concerns is the “one and done” mechanism [8]. Because only one adenoma is needed to score a detection in ADR calculation, this could incentivize the removal of a single adenoma followed by less careful examination of the remaining colon. Indeed, one clear example of “one and done” behavior has been documented [9], in which veterans undergoing colonoscopy under contract by private practice colonoscopists had a similar ADR to veterans undergoing screening colonoscopy within the VA hospital system, but adenomas per colonoscopy (APC) was lower for the private practice doctors compared to the VA hospital physicians [9]. This observation supports a move from ADR to adenomas per colonoscopy (APC) as the primary quality measure, and the possibility of this move was discussed in detail in the most recent revision of the ACG/ASGE quality recommendations [3]. There remains some concern about where thresholds for acceptable performance should be set for APC, and some concern that APC would incentivize placement of individual polyps into separate containers for pathology, thereby increasing pathology costs [3]. This risk could be mitigated by a mandate of photography for individual polyps, but many practicing colonoscopists do not have the technology available to take and store photographs of each individual polyp detected. Further, other than the study cited above, a large number of studies show that ADR and APC have an excellent correlation [10–12]. Also, ADR is now validated as a proven predictor of colorectal cancer prevention [13, 14], while APC is not. Finally, it is quite possible that the behavior demonstrated in the single documented example of “one and done” gaming reflects the reimbursement system rather than ADR, because doctors in the U.S. (including in the VA colonoscopy system contracting with external sources) receive payment for the first polyp removed but not for the resection of subsequent polyps.

We suggest that confining ADR to screening examinations only creates a second potential avenue for gaming ADR, which we term “indication gaming.” The potential for indication gaming arises when patients have multiple indications for colonoscopy. For example, a 60-year-old patient with no prior colonoscopy who presents for evaluation of rectal bleeding or a change in bowel habits could be considered to have both the symptom and screening as indications. The problem of gaming arises because the procedure report is generated after the procedure has been completed, at which time the colonoscopist often knows with high probability whether a conventional adenoma has been detected. The current system of measuring ADR incentivizes the doctor to call the examination a screening examination if an adenoma had been detected, and a diagnostic examination using the patient’s symptom(s) as the indication if no adenoma was detected. Such behavior, to the extent it occurs in practice, constitutes gaming.

In the first portion of this study, we retrospectively assessed our quality database to evaluate whether determining an overall ADR based on screening, surveillance, and diagnostic examinations significantly impacts conclusions about endoscopist performance compared to the standard methodology of confining ADR to screening-only examinations. In the second part of the study, we prospectively assessed the feasibility of gaming the screening-only ADR by having an expert colonoscopist shift the indication in patients with both screening and diagnostic indications, according to his impression of whether a conventional adenoma had been detected during the procedure.

Methods
We retrospectively evaluated the screening, surveillance, and diagnostic examination ADRs of 15 endoscopists who are currently practicing in our unit and who have at least 200 examinations in persons aged ≥50 years between July 2010 and January 2015. In our unit, we maintain a quality database that allows calculation of ADR for all physicians and for all indications. For this database, we used the indication actually listed on the procedural report to designate the procedure as “screening,” “surveillance,” or “diagnostic.” In the small number of patients where the indication listed a diagnostic indication plus either screening or surveillance we used the indication listed first for that procedure. Patients with inflammatory bowel disease, polyp syndromes, Lynch syndrome, and high risk family histories (multiple first degree relatives with colorectal cancer, or a single first degree relative with colorectal cancer at <60 years) were excluded. Patients in the surveillance group had prior colorectal polyps but not cancer. We present the ADRs for screening, surveillance, and diagnostic examinations and the overall ADR for all three types of examination combined, with the emphasis on the absolute differences between screening ADR and overall ADR for each physician.

For the second part of the study all patients presenting for screening colonoscopy by the senior author (D.K.R.) were prospectively assessed. Screening was considered a potential indication if the patient had no prior colonoscopy and was aged ≥50 years. If patients referred for screening had symptoms that they had raised with their primary care physician and which were included on the referral sheet, and which they acknowledged in the pre-procedure interview, or for which they initiated a discussion during the pre-procedure interview and consent process, then they were considered to also have symptoms that were a potential diagnostic indication for the examination. When the procedure was over, the colonoscopist considered which indication would be used based on his impression of whether a conventional adenoma had been identified. Patients with both screening and symptomatic indications were given the symptomatic indication it if the colonoscopist believed no conventional adenoma had been detected. If the
Table 1 shows the numbers of examinations performed by each of 15 physicians with more than 200 examinations on persons aged ≥50 years during the study interval. The surveillance ADR was higher than the screening ADR for all physicians except one (physician 15; Fig. 1). The screening ADR was higher than the diagnostic ADR for all but one physician (physician 1; Fig. 1). The mean absolute difference between the screening ADR and the overall ADR for the three category indications combined was 2.6 percentage points (range 0–6.9 percentage points). The mean absolute difference between the surveillance ADR and the overall ADR was 8.1 percentage points (range 0.3–11.9 percentage points). The mean absolute difference between the diagnostic ADR

colonoscopy believed that a patient with both screening and symptom indications did have a conventional adenoma detected, the patient was given screening as the indication. The colonoscopy had extensive experience in differentiation of conventional adenomas from serrated class lesions [15–17].

We had no data to base a sample size on, so we chose a convenience sample of approximately 100 consecutive patients with either a screening indication, or combined screening and diagnostic indications. The final sample was 118 patients.

Permission to conduct the study was granted by the Institutional Review Board at Indiana University Purdue University Indianapolis on August 26, 2016.
and the overall ADR was 10.9 percentage points (range 3.1 – 17.7 percentage points).

Using the older recommendation for acceptable ADR threshold for a mixed gender population of 20% [1, 3], all 15 physicians were above the recommended threshold using either the screening ADR or the overall ADR (Table 1). Using the newer recommended minimum threshold of 25% for a mixed gender population, one physician (physician 8; Table 1) moved from just below the recommended threshold with a screening ADR of 24.6% to an ADR just above the threshold at 26.5% using the combined indication ADR measurement. All of the remaining physicians were above the recommended threshold with both the screening ADR and the combined overall ADR (Table 1).

Table 2 shows the mean ages and the percent of males for the patients of each physician, according to the category of indication.

To evaluate the feasibility of “indication gaming” on ADR, a single endoscopist (D. K. R.) identified 118 consecutive colonoscopies that were either straightforward first time screening examinations (n = 93), or where patients were presenting at age ≥50 years for first time examinations but who also had symptoms designated on the referral sheet or who offered symptoms to the colonoscopist during the pre-procedure interview (n = 25). The ADR for the 93 straightforward screening examinations was 48.4% (45 of 93 patients had ≥1 conventional adenoma). Of the 25 patients with both screening and diagnostic indications, 14 had ≥1 conventional adenoma, and the presence or absence of a conventional adenoma was correctly predicted by the colonoscopist after the procedure in all 25 cases. By designating all 14 patients with predicted adenomas as “screening” for the procedure indication and designating the 11 patients without a predicted adenoma as having a diagnostic indication, the endoscopist increased the calculated screening ADR from 48.4% to 55.1% (after “indication gaming,” there were 59 of 107 screening patients with ≥1 conventional adenoma).

Table 2: Mean age and percentage of male patients in the study by physician and category of indication for colonoscopy.

<table>
<thead>
<tr>
<th>Physician no.</th>
<th>Diagnostic</th>
<th></th>
<th>Screening</th>
<th></th>
<th>Surveillance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean age, years</td>
<td>% Male</td>
<td>Mean age, years</td>
<td>% Male</td>
<td>Mean age, years</td>
<td>% Male</td>
</tr>
<tr>
<td>1</td>
<td>61.7</td>
<td>38.4%</td>
<td>57.6</td>
<td>42.8%</td>
<td>62.7</td>
<td>45.5%</td>
</tr>
<tr>
<td>2</td>
<td>64.6</td>
<td>39.9%</td>
<td>58.7</td>
<td>48.2%</td>
<td>62.7</td>
<td>56.4%</td>
</tr>
<tr>
<td>3</td>
<td>62.4</td>
<td>43.4%</td>
<td>58.7</td>
<td>43.9%</td>
<td>64</td>
<td>46%</td>
</tr>
<tr>
<td>4</td>
<td>66.8</td>
<td>31.6%</td>
<td>56.4</td>
<td>24.6%</td>
<td>64.3</td>
<td>37.1%</td>
</tr>
<tr>
<td>5</td>
<td>64.4</td>
<td>31.3%</td>
<td>58.6</td>
<td>35%</td>
<td>63</td>
<td>56.9%</td>
</tr>
<tr>
<td>6</td>
<td>64.7</td>
<td>43.4%</td>
<td>59.1</td>
<td>54.6%</td>
<td>63.9</td>
<td>58.2%</td>
</tr>
<tr>
<td>7</td>
<td>63.3</td>
<td>40.9%</td>
<td>58.9</td>
<td>50.5%</td>
<td>63.4</td>
<td>53%</td>
</tr>
<tr>
<td>8</td>
<td>61.9</td>
<td>41.2%</td>
<td>58.7</td>
<td>44.9%</td>
<td>62.2</td>
<td>60.7%</td>
</tr>
<tr>
<td>9</td>
<td>65.3</td>
<td>34.2%</td>
<td>59.2</td>
<td>48.4%</td>
<td>64</td>
<td>58.4%</td>
</tr>
<tr>
<td>10</td>
<td>62.2</td>
<td>26.2%</td>
<td>57</td>
<td>42.7%</td>
<td>63.3</td>
<td>55.2%</td>
</tr>
<tr>
<td>11</td>
<td>63.8</td>
<td>38.6%</td>
<td>60.6</td>
<td>49.9%</td>
<td>65.8</td>
<td>52.6%</td>
</tr>
<tr>
<td>12</td>
<td>64.2</td>
<td>33.8%</td>
<td>58.6</td>
<td>53.8%</td>
<td>64.3</td>
<td>45.8%</td>
</tr>
<tr>
<td>13</td>
<td>62.3</td>
<td>44.2%</td>
<td>60.8</td>
<td>51.9%</td>
<td>65.6</td>
<td>43.8%</td>
</tr>
<tr>
<td>14</td>
<td>63.7</td>
<td>39.7%</td>
<td>59.5</td>
<td>49.1%</td>
<td>61.9</td>
<td>55.9%</td>
</tr>
<tr>
<td>15</td>
<td>62.3</td>
<td>35.8%</td>
<td>59.7</td>
<td>52.9%</td>
<td>63.7</td>
<td>62.6%</td>
</tr>
</tbody>
</table>
Discussion

In this report we demonstrate that the screening ADR and the overall ADR (combining screening, surveillance and screening indications) are very similar, and that the performance of an individual physician is unlikely to be affected by the use of screening vs. overall ADR.

As in previous studies, we found that surveillance ADRs were higher than screening ADRs, which were higher than diagnostic ADRs [12, 18]. Because surveillance ADRs are higher than screening ADRs, and screening ADRs are higher than diagnostic ADRs, the overall ADR which encompasses all three types of examination tends to be closer to the screening ADR than the surveillance or diagnostic ADR. Our results suggest that narrowing the ADR to screening indications only is unnecessary. Expanding ADR back to include screening, surveillance, and diagnostic examinations (with exclusion only of patients under age 50 and those with inflammatory bowel disease, polyp syndromes, or other inherited colorectal cancer syndromes) appears feasible. Re-expanding the target for ADR to include all three types of examinations would also make measurement of ADR easier for doctors and would give narrower confidence intervals around measurements.

The reasons why the diagnostic ADR runs below the screening and surveillance ADRs are not currently clear, and are a target for further investigation. Lower mean age is not an explanation for the differences, as shown in Table 2. Possibly, patients with diagnostic examinations are generally more likely to have undergone prior negative colonoscopies, and thus are selected for further negative examinations.

We also showed that “indication gaming” of the screening ADR is feasible. Since even the potential for gaming undermines a quality indicator, this finding supports simplifying the ADR to a measurement that includes most colonoscopies in persons aged >50 years.

Limitations of this study include that it was performed at a single center. Of the 15 study physicians, 11 had overall and screening ADRs above 30%. Thus, the results may not be generalizable, particularly in centers with doctors having relatively lower ADRs. Evaluation of these relationships at other centers is important. Some physicians might have high fractions of colonoscopies with diagnostic indications, which could be problematic when applying an overall ADR. In our study the percent of examinations classified as diagnostic ranged from 19% to 31% for 13 of the 15 physicians (Table 1), with 2 physicians who were outliers at 15% and 50%, respectively. This fraction of diagnostic examinations might need to be monitored in some practices, and further study is warranted.

In summary, we demonstrated that in our practice a simplified overall ADR encompassing screening, surveillance, and diagnostic examinations would be just as effective as a screening-only ADR for establishing the mucosal inspection skills of colonoscopists. Further, we showed the feasibility of gaming a screening-only ADR. We recommend that others evaluate this issue in their quality databases. If confirmed by others, we suggest that a return to the original definition of ADR [3] be considered when recommendations for quality indicators are updated.

Acknowledgment

This work was supported by a gift to the Indiana University Foundation by Scott and Kay Schurz, and their children of Bloomington, Indiana, USA.

Competing interests

D. K. Rex is a consultant for Olympus Corp. He has consultancy and research support links with Boston Scientific and research support links with Endochoice and EndoAid.

References


