Comparison of Quality Performance Metrics in Screening and Surveillance Colonoscopy: A Single-Center Experience

James Stephen Love (✉ jlove25@uic.edu)
UIC: University of Illinois Chicago
https://orcid.org/0000-0002-5727-8121

Michael Siegel
UIC: University of Illinois Chicago

Meredith Yellen
UIC: University of Illinois Chicago

Jeffrey Rebhun
Oregon Health & Science University University Hospital: Oregon Health & Science University Hospital

Asim Shuja
UIC: University of Illinois Chicago

Research Article

Keywords: Quality improvement, benchmark, colonoscopy, surveillance, Adenoma detection rate

Posted Date: January 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2046463/v1

License: ☕️ 📧 This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background and aims:

Current guidelines for screening colonoscopies recommend a minimum adenoma detection rate (ADR) of 25%. There are no established benchmarks for surveillance colonoscopies and data surrounding the utility of other quality metrics in this setting is limited. We aimed to define the relationship between ADR and alternative quality measures in the setting of screening and surveillance colonoscopies and determine whether validated screening quality benchmarks can be extrapolated to surveillance procedures.

Methods:

A retrospective review was conducted of adults who underwent outpatient screening and surveillance colonoscopies at a tertiary health center. ADR, adenomas per colonoscopy (APC), adenomas per positive participant (APP), polyp detection rate (PDR), right-sided polyp detection rate (RSP), and withdrawal times (CWT) were analyzed for screening and surveillance colonoscopies. Normality tests were performed for each continuous variable. Kruskal-Wallis tests were performed to analyze categorical and continuous outcomes. Spearman rank correlations were performed to compare surrogate outcomes to ADR.

Results:

In total, 2646 procedures, 1884 screening and 762 surveillance, were analyzed. Surveillance ADR (CADR) was significantly higher than screening ADR (SADR) (65.6% ± 0.02 v. 47.0% ± 0.01%, p<0.001). All alternate quality measures except CWT were significantly higher in surveillance procedures. Among surveillance procedures, there was a strong correlation between CADR and PDR (r=0.956, p<0.01), as well as RSP (r=0.771, p=0.003); correlations between CADR and other alternate quality measures were not significant.

Conclusion:

Colonoscopy quality measures were significantly higher in surveillance procedures when compared to screening procedures despite similar CWT. Higher benchmarks must be established to ensure quality surveillance colonoscopies.

Introduction

Colorectal cancer (CRC) is the third most common cause of cancer in the United States with an estimated cost of over $17 billion per year [1–2]. Screening colonoscopies allow for early detection and removal of precancerous colonic lesions, and reduce the incidence of colorectal cancer [1, 3]. For these reasons, colonoscopy is heralded as the gold-standard CRC screening tool. In order to ensure high-quality colonoscopy performance, several quality indicators have been introduced. Adenoma detection rate (ADR), defined as the proportion of screening colonoscopies performed that detect at least one histological colorectal adenoma or adenocarcinoma, is the most widely studied and validated quality indicator for screening colonoscopy [4]. Initially, the ADR benchmark of > 25% was considered adequate;
however, recent trends in screening colonoscopies have shown significant increases in average ADR, estimated to be approximately 40%, suggesting the need to raise the ADR benchmark [4–5].

There are several factors that may impact ADR for routine screening colonoscopies, such as endoscopist experience, patient gender, day of the week, and bowel-prep adequacy [6–7]. Consequently, adenoma detection is endoscopist-dependent, with considerable variability between endoscopists [8]. Further complicating this is the observation that high ADR does not always correlate with lower adenoma miss rates (AMR), suggesting that ADR as a sole indicator of colonoscopy quality performance is likely inadequate [8–10]. In an attempt to reduce variation and provide additional information on endoscopist performance, several other quality indicators have been introduced, including adenomas per colonoscopy (APC), adenomas per positive participant (APP), polyp detection rate (PDR), colonoscopy withdrawal time (CWT), and right sided polyp detection (RSP) [4, 8, 11–13].

Although these quality parameters have demonstrated utility in distinguish high-performing endoscopists beyond the standard ADR metric in screening colonoscopies, research regarding their role in surveillance procedures is limited and there are no established benchmarks to ensure quality colonoscopy performance in this setting. This study aims to define the relationship between ADR and alternative quality indicators (APC, APP, PDR, CWT, RSP) in the setting of screening and surveillance colonoscopies and determine whether validated screening quality benchmarks can be extrapolated to surveillance procedures.

Materials And Methods

Study design

This was a retrospective observational study of all adult patients who underwent outpatient screening and surveillance colonoscopies from January 2015 to April 2020. Procedures were completed at a single university tertiary health center by board-certified gastroenterologists. This study was approved by the Institutional Review Board at the University of Illinois Hospital and Health Sciences System in March 2020.

Study Population

The study population was composed of 2646 patients undergoing outpatient screening and surveillance colonoscopies from January 2015 to April 2020. Patients with a history of inflammatory bowel disease and colorectal cancer were excluded from the study. Procedures with failed cecal intubation, inadequate bowel preparation, and diagnostic intentions were also excluded.

Definition of variables and outcomes:

ADR, APC, APP, PDR, RSP, and CWT were analyzed for screening and surveillance colonoscopies. Screening colonoscopy was defined as a colonoscopy performed on an asymptomatic patient without history of colonoscopy or history of adenomatous polyps on previous colonoscopy. Surveillance
Colonoscopy was defined as a colonoscopy performed on a patient with history of adenomatous polyps on previous screening colonoscopy or CRC. ADR was defined as the number of colonoscopies performed that detect at least one histological colorectal adenoma or adenocarcinoma divided by the number of colonoscopies. APC was defined as the number of detected adenomas divided by the total number of screening colonoscopies. APP was defined as the number of detected adenomas divided by the number of screening colonoscopies in which one or more adenomas are detected. PDR was defined as the number of colonoscopies in which at least 1 polyp was detected divided by the total number of colonoscopies. RSP was defined as the number of colonoscopies in which at least 1 polyp was detected in the right colon divided by the total number of colonoscopies. CWT was defined as the time spent visualizing the colon from cecal intubation to withdrawal of the scope from the anus. To better estimate the time spent visualizing the colon, withdrawal times were corrected for by the total number of polyps resected respective of the individual procedure. Bowel preparation was scored using the Boston Bowel Preparation Scale (BBPS). A procedure with a score of less than 6 was deemed as having inadequate prep.

Statistical analysis:

Categorical data was presented as percentages and continuous data presented as mean ± standard deviation. Chi-square analysis was performed to compare categorical variables. Normality tests were performed for each continuous variable and Mann Whitney U or Kruskal-Wallis tests were performed to analyze continuous outcomes as appropriate. A two-sided p-value of less than 0.05 was used to determine statistical significance. Spearman rank correlations were performed to compare surrogate performance outcomes to ADR. A strong correlation was defined by coefficient values between 0.7 and 1. Statistical analyses were performed using IBM SPSS Statistics for Macintosh, Version27.0 (IBM Corp., Armonk, NY).

Results

A total of 2646 procedures were included in our analysis. The average patient age for screening colonoscopies was 57.4 years (± 7.9 years) and 44.5% were men (n = 838); the average patient age for surveillance colonoscopies was 62.2 years (± 8.7 years) and 45.4% were men (n = 346). Patient demographics are described in Fig. 1. Screening colonoscopies accounted for 71.2% of included procedures (n = 1884), while surveillance colonoscopies accounted for the remaining 28.8% (n = 762).

Quality indicators for screening and surveillance colonoscopies are summarized in Table 1. Surveillance ADR (CADR) was significantly higher than screening ADR (SADR) (65.6%+0.02 v. 47.0%+0.01%, p < 0.001). All alternate quality measures including ADR, APC, PDR, APP, and RSP were significantly higher in surveillance procedures, with the exception of CWT. Among surveillance procedures, there was a strong correlation between CADR and PDR (r = 0.956, p < 0.01), as well as RSP (r = 0.771, p = 0.003); correlations between CADR and other alternate quality measures, including APP and APC, were not statistically significant (Table 2).
Table 1
Association of quality parameters between screening and surveillance procedures.

<table>
<thead>
<tr>
<th>Quality Parameter</th>
<th>Screening (n = 1884)</th>
<th>Surveillance (n = 762)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADR</td>
<td>47.0% ± 0.01</td>
<td>65.6% ± 0.02</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>APC</td>
<td>1.60 ± 0.04</td>
<td>1.81 ± 0.07</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>APP</td>
<td>2.07 ± 0.06</td>
<td>2.30 ± 0.08</td>
<td>0.012</td>
</tr>
<tr>
<td>PDR</td>
<td>64.7% ± 0.01</td>
<td>77.7% ± 0.02</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>RSP</td>
<td>45.3% ± 0.01</td>
<td>63.3% ± 0.02</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>CWT</td>
<td>15.8 ± 9.2</td>
<td>16.5 ± 9.7</td>
<td>0.079</td>
</tr>
</tbody>
</table>


Table 2
Quality parameters and their association with ADR in surveillance procedures.

<table>
<thead>
<tr>
<th>Quality Parameter</th>
<th>Spearman Rank Correlation ((r))</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>0.699</td>
<td>0.011</td>
</tr>
<tr>
<td>APP</td>
<td>0.21</td>
<td>0.513</td>
</tr>
<tr>
<td>PDR</td>
<td>0.956</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>RSP</td>
<td>0.771</td>
<td>0.003</td>
</tr>
<tr>
<td>CWT</td>
<td>-0.049</td>
<td>0.88</td>
</tr>
</tbody>
</table>


Discussion

In this study, we evaluated quality performance metrics in outpatient screening and surveillance colonoscopies at a single-center tertiary health center. Overall, we found that all quality measures are significantly higher in surveillance procedures when compared to screening procedures despite similar CWT, suggesting that separate ADR targets for surveillance and screening procedures be established to ensure quality performance. This is inconsistent with previously reported data wherein similar rates of ADR and other quality metrics are found between screening and surveillance procedures [14].

Our data suggests that colonoscopy quality measures are significantly higher in surveillance procedures when compared to screening procedures despite similar CWT. This is inconsistent with previous similar studies, wherein CADR and SADR were not significantly different, and far exceeds previously reported
CADR rates [14–16]. The SADR among our cohort was approximately 47%, which exceeds the recommended minimum benchmark of 25%. As mentioned, while S-ADR can vary widely depending on certain patient populations and practice settings, according to a large national quality benchmarking registry, recent trends in screening colonoscopy performance suggest that the SADR achieved in most practices is closer to 35–40%, which is similar to the SADR observed in our tertiary health center [5, 17]. However, as the entry age for CRC screening has recently been lowered to include those aged 45 years and older, and while alternatives such as the fecal immunochemical test (FIT) are becoming increasingly popular, it is reasonable to suspect that there will be larger demand for surveillance colonoscopies compared to screening colonoscopies in the near future [18–19]. In light of this, we should anticipate a larger volume of surveillance procedures and establish appropriate benchmarks to ensure quality surveillance colonoscopy performance and to improve detection rates of interval CRC.

In addition, we aimed to directly compare quality measures across surveillance colonoscopies and explore their correlations. In a previous study, it was found that all quality parameters except CWT correlate strongly with ADR in screening colonoscopies [6]. Interestingly, in the setting of surveillance colonoscopies, we found that only PDR and RSP are strongly associated with CADR. This novel finding suggests that APP and APC may provide additional information regarding surveillance endoscopist performance, though further studies are needed to strengthen this association.

Few studies have directly compared the utility of ADR and other quality indicators in screening and surveillance procedures, and there are no established benchmarks for quality indicators for surveillance procedures [14, 16]. Our findings directly aid in closing this gap in our scientific literature and are strengthened by a large sample size of over 2600 colonoscopies performed by 29 endoscopists at a major academic tertiary center with a diverse patient population. However, this study is not devoid of limitations. This is a single-center study, and thus it is difficult to assess the generalizability and implications of these results at different practices. Additionally, endoscopists were not blinded during this study and each endoscopist was aware of each patient's history and reason for colonoscopy. The heightened awareness of reducing interval CRC incidence during surveillance procedures may have resulted in improved colonoscopy performance and a drastically higher CADR, as seen in this study; though, the mechanism of this is unclear as there was no significant difference in CWT between surveillance and screening procedures.

**Conclusion**

This is the first study demonstrating the need for higher quality benchmarks for surveillance colonoscopies relative to screening colonoscopies. This study confirms recent trends in screening colonoscopies, echoing the need for a higher SADR benchmark and demonstrating a positive correlation between SADR and all other quality metrics in screening procedures. In this study, all quality indicators were significantly higher in surveillance procedures despite similar CWTs. Furthermore, among surveillance procedures, CADR did not correlate with APP and APC, suggesting these quality indicators may provide additional information regarding surveillance endoscopist performance. Despite these
encouraging findings, we recommend similar prospective studies be conducted in order to establish definitive quality benchmarks for surveillance colonoscopies and to ensure endoscopist quality performance in a larger variety of settings.

**Abbreviations**

CRC  
colorectal cancer  
AMR  
adenoma miss rate  
ADR  
adenoma detection rate  
APC  
adenomas per colonoscopy  
APC  
adenomas per positive participant  
PDR  
polyptidection rate  
RSP  
right-sided polyptidection rate  
CWT  
colonoscopy withdrawal time  
CADR  
surveillance adenoma detection rate  
SADR  
screening adenoma detection rate  
BBPS  
Boston Bowel Preparation Scale  
FIT  
fecal immunohistochemical test

**Declarations**

**Ethics approval and consent to participate:**

This study was approved by the Institutional Review Board at the University of Illinois Hospital and Health Sciences System in March 2020.

**Consent for publication**

Not applicable
Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Funding:

There are funding sources to declare.

Authors' contributions:

JL, MY, JR, and AS collected, analyzed, and interpreted the patient data. JR performed the formal statistical analyses. JL wrote the manuscript. MS, MY, JR, and AS assisted in the editing and revision process. All authors read and approved the final manuscript.

Acknowledgements:

Not applicable.

References


Figures
<table>
<thead>
<tr>
<th>Variable</th>
<th>Screening</th>
<th>Surveillance</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1046</td>
<td>416</td>
<td>1462</td>
</tr>
<tr>
<td>Male</td>
<td>838</td>
<td>346</td>
<td>1184</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>57.4 +/- 7.9</td>
<td>62.2 +/- 8.7</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>1044</td>
<td>432</td>
<td>1476</td>
</tr>
<tr>
<td>Caucasian</td>
<td>316</td>
<td>155</td>
<td>471</td>
</tr>
<tr>
<td>Latino</td>
<td>344</td>
<td>131</td>
<td>475</td>
</tr>
<tr>
<td>Asian</td>
<td>93</td>
<td>15</td>
<td>108</td>
</tr>
<tr>
<td>Other</td>
<td>58</td>
<td>21</td>
<td>79</td>
</tr>
<tr>
<td>Unknown</td>
<td>29</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td><strong>Bowel Prep</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>1164</td>
<td>570</td>
<td>1734</td>
</tr>
<tr>
<td>Excellent</td>
<td>720</td>
<td>192</td>
<td>912</td>
</tr>
</tbody>
</table>

AA - African American.

**Figure 1**

Baseline patient characteristics of study population.