REVIEW

see related editorial on page 1725

CME

What Level of Bowel Prep Quality Requires Early Repeat Colonoscopy: Systematic Review and Meta-Analysis of the Impact of Preparation Quality on Adenoma Detection Rate

Brian T. Clark, MD[1], Tarun Rustagi, MD[1] and Loren Laine, MD[1,2]

OBJECTIVES: **Current guidelines recommend early repeat colonoscopy when bowel preparation quality is inadequate, defined as inability to detect polyps >5 mm, but no data link specific bowel preparation categories or scores to this definition. Nevertheless, most physicians use a shortened screening/surveillance interval in patients with intermediate-quality preparation. We determined whether different levels of bowel preparation quality are associated with differences in adenoma detection rates (ADRs: proportion of colonoscopies with ≥1 adenoma) to help guide decisions regarding early repeat colonoscopy—with primary focus on intermediate-quality preparation.**

METHODS: **MEDLINE and Embase were searched for studies with adenoma or polyp detection rate stratified by bowel preparation quality. Preparation quality definitions were standardized on the basis of Aronchick definitions (excellent/good/fair/poor/insufficient), and primary analyses of ADR trichotomized bowel preparation quality: high quality (excellent/good), intermediate quality (fair), and low quality (poor/insufficient). Dichotomized analyses of adequate (excellent/good/fair) vs. inadequate (poor/insufficient) were also performed.**

RESULTS: **Eleven studies met the inclusion criteria. The primary analysis, ADR with intermediate- vs. high-quality preparation, showed an odds ratio (OR) of 0.94 (0.80–1.10) and absolute risk difference of −1% (−3%, 2%). ADRs were significantly higher with both intermediate-quality and high-quality preparation vs. low-quality preparation: OR = 1.39 (1.08–1.79) and 1.41 (1.21–1.64), with absolute risk increases of 5% for both. ADR and advanced ADR were significantly higher with adequate vs. inadequate preparation: OR = 1.30 (1.19–1.42) and 1.30 (1.02–1.67). Studies did not report other relevant outcomes such as total adenomas per colonoscopy.**

CONCLUSIONS: **ADR is not significantly different with intermediate-quality vs. high-quality bowel preparation. Our results confirm the need for early repeat colonoscopy with low-quality bowel preparation, but suggest that patients with intermediate/fair preparation quality may be followed up at standard guideline-recommended surveillance intervals without significantly affecting quality as measured by ADR.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

Am J Gastroenterol 2014; 109:1714–1723; doi:10.1038/ajg.2014.232; published online 19 August 2014

INTRODUCTION

Colorectal cancer is the second leading cause of cancer deaths in the United States, accounting for >50,000 deaths annually (1). Given the ability to not only identify but also remove precursor lesions, colonoscopy has emerged as the gold standard for screening and surveillance of colorectal neoplasia, and has been shown to reduce mortality from colorectal cancer (2–5). The identification and removal of adenomatous polyps is critical to the success of screening and surveillance colonoscopy, and the adenoma detection rate (ADR), which correlates with the development of interval colorectal cancer, is widely used as a measure of colonoscopy quality (6,7). The quality of bowel preparation determines

[1]Section of Digestive Diseases, Yale University School of Medicine, New Haven, Connecticut, USA; [2]VA Connecticut Healthcare System, West Haven, Connecticut, USA.
Correspondence: Loren Laine, MD, Section of Digestive Diseases, Yale University School of Medicine, P.O. Box 208019, New Haven, Connecticut 06520-8019, USA.
E-mail: loren.laine@yale.edu
Received 9 February 2014; accepted 3 June 2014

the endoscopist's ability to visualize the colonic mucosa, and it appears to be a major factor affecting the ability to detect adenomas and carry out a high-quality examination (2,6,8–10).

Current guidelines suggest early repeat colonoscopy when bowel preparation quality for a screening or surveillance colonoscopy is inadequate, defined as the inability to identify lesions >5 mm (11,12). However, no data are available to determine how to judge whether the bowel preparation is adequate to identify lesions >5 mm, thus limiting the ability to specify the appropriate time interval for a repeat colonoscopic examination in any patient without an excellent bowel preparation. In addition, measures of bowel preparation quality used in practice are inconsistent and not standardized, leading to significant variability, both in how bowel preparation quality is scored and how physicians determine timing of surveillance intervals based on preparation quality (13).

In clinical practice, patients with high-quality preparation (e.g., excellent or good) generally return at standard guideline-recommended intervals and those with low-quality preparation (e.g., poor or insufficient) return earlier. However, uncertainty exists over the appropriate interval for patients with preparation quality between high and low, often defined as "fair" or "intermediate" (13–15). Recent studies report that physicians in practice use a shortened interval in the majority of patients with intermediate-quality preparation (13,14), despite a lack of evidence indicating that ADR or other clinical outcomes are worse with intermediate-quality than with high-quality preparation.

The goal of this systematic review and meta-analysis is to determine whether different levels of bowel preparation quality are associated with differences in ADR and thereby assist in decisions regarding the time interval for repeat screening or surveillance colonoscopy. Our primary analysis evaluates whether intermediate-quality bowel preparation is associated with an ADR that is inferior to high-quality bowel preparation. Our secondary aims are to compare ADR in patients with (i) intermediate-quality vs. low-quality preparation, (ii) high-quality vs. low-quality preparation, (iii) excellent vs. good preparation, and (iv) adequate vs. inadequate preparation.

METHODS

The MOOSE (Meta-analysis of Observational Studies in Epidemiology) statement and guidelines were consulted during the stages of design, analysis, and reporting of this meta-analysis (16). The search strategy, study inclusion and exclusion criteria, primary and secondary outcomes, and analyses were defined on an *a priori* basis and are described in this section.

Search strategy/data extraction

Systematic searches of MEDLINE and Embase databases were performed from inception through October 2013 using the OvidSP interface. Search terms included both MeSH and non-MeSH terms relating to colonoscopy, bowel preparation, and colonic polyps or adenomas. The search strategy, shown in **Supplementary Figure S1** online, was engineered by two authors (B.T.C. and L.L.) with the assistance of a Yale University medical librarian. Searches had no language restriction but were restricted to human and adult studies. Translation of potentially relevant non-English articles was performed by a reviewer fluent in that language if no English translation of the article was discovered. In addition, a recursive search of the reference sections of selected studies, practice guidelines, and pertinent review articles was performed to identify other potentially relevant articles.

Review of titles/abstracts, full review of potentially relevant studies to identify those meeting inclusion criteria, and data extraction using a standardized collection form were performed independently by two authors (B.T.C. and T.R.). Titles/abstracts considered potentially relevant by either reviewer were retrieved for review of the full article. The lists of full articles meeting inclusion criteria from the two reviewers were compared and any disagreements were resolved by discussion and consensus, with the senior author (L.L.) serving as the final arbiter if consensus was not achieved. The same process was used for resolution of any disagreements upon comparison of the data extraction forms from the two reviewers. Authors were not contacted directly for additional unpublished data.

Study selection and assessment of methodologic quality

Studies eligible for inclusion met the following criteria: (i) study design: experimental or observational; (ii) study population: patients undergoing colonoscopic evaluation; (iii) "intervention": bowel preparation quality was defined and reported, allowing data extraction using our predefined bowel preparation quality categories (discussed below); and (iv) outcome measure: adenoma or polyp detection rate was reported (in the form of raw numbers or odds ratios (ORs)) for at least two of the desired strata of bowel preparation quality. Studies not meeting these criteria were excluded.

We evaluated the QUADAS-2 tool and the Newcastle–Ottawa scale for assessing the quality and internal validity of the nonrandomized component studies in our meta-analysis; however, these tools were not directly applicable to studies included in our systematic review (17,18). We therefore adapted a quality assessment tool directed at our component studies using these tools as a reference (**Supplementary Figure S2**). This tool evaluated the representativeness of the subject population (consecutive enrollment of subjects, screening/surveillance colonoscopies), ascertainment of exposure (use of a validated quantitative or ordinal scale for bowel preparation quality), full accounting of all subjects, comparability of cohorts (use of regression analysis to control for potential confounders), and assessment of the outcome (use of ADR rather than polyp detection rate (PDR)). On the basis of the Newcastle–Ottawa scale, a maximum of 2 points were assigned for the item related to control of potentially confounding variables, whereas a maximum of 1 point was assigned for the other 5 items. *A priori*, we defined a high-quality study as one meeting a threshold of ≥5 of 7 possible points.

Definition of bowel preparation quality. The reported definitions of bowel preparation quality from component studies were standardized based on the Aronchick (19) definitions.

Excellent quality is defined as a small volume of clear liquid with >95% of mucosal surface seen; good quality is defined as a large volume of clear liquid but >90% of mucosal surface seen; fair quality is defined as some semisolid stool that could be suctioned or washed away but >90% of mucosal surface seen; poor quality indicates semisolid stool that could not be suctioned or washed away with <90% of mucosal surface seen; and insufficient quality indicates that fecal material could not be cleared with repeat preparation required. We chose the Aronchick scale for standardization because it is the most widely used and studied scale, and because of the paucity of large studies using other well-validated scales such as the Boston (15,20) or Ottawa (21) bowel preparation scales.

After standardization of bowel preparation scores from component studies, we trichotomized groups for primary analyses: (i) high quality (excellent/good preparation); (ii) intermediate quality (fair preparation); and (iii) low quality (poor/insufficient preparation). For additional analyses, excellent, good, and fair preparations were defined as adequate, whereas poor and insufficient preparations were defined as inadequate. In this analysis, a Boston Bowel Preparation score (BBPS) of ≥5 was considered adequate (excellent/good/fair) and BBPS of <5 was considered inadequate (poor/inadequate), as suggested by the validation study (20).

Definition of ADR. ADR was defined within each strata of bowel preparation quality as the proportion of colonoscopies with at least one adenoma detected. ORs and 95% confidence intervals (CIs) from component studies were preferentially used for the comparisons of ADR. If ORs were not directly provided, then ADRs were calculated from the raw data provided by the study, from which ORs and standard errors were calculated. If only PDR was provided, ADR was estimated by multiplying PDR×0.645 (22,23). We also performed separate exploratory analyses of advanced ADR, for which component studies were required to provide data on the proportion of colonoscopies with at least one advanced adenoma (defined by the presence of significant villous features, high-grade dysplasia or invasive features, and/or size ≥1 cm).

Data analysis

The effect of bowel preparation quality on ADR was calculated using a pooled estimate of OR with 95% CIs. Given that some component studies provided only OR without raw data, we used the generic inverse variance method to calculate the pooled OR (24). Heterogeneity was calculated using the χ^2 test and I^2 statistic. When heterogeneity existed among component studies ($P<0.10$), a random-effects model was used to calculate the pooled OR; otherwise, a fixed-effect model was used. Our primary analysis compared ADR for patients with intermediate-quality preparation vs. high-quality preparation. Our secondary analyses evaluated ADR for (i) intermediate-quality vs. low-quality preparation; (ii) high-quality vs. low-quality preparation; (iii) excellent-quality vs. good-quality preparation; and (iv) adequate vs. inadequate preparation. Subsequently, we compared the advanced ADR for each of the aforementioned comparisons.

We prespecified the following subgroup analyses to identify the influence of different study characteristics on our primary outcome: (i) studies that reported ADR vs. PDR and (ii) studies that evaluated only screening or surveillance colonoscopy vs. those that evaluated any colonoscopy indication. Treatment effect by subgroup interaction was assessed by calculating the heterogeneity between subgroups with significant heterogeneity defined as $P<0.05$ using the χ^2 test. We also prespecified sensitivity analyses of our primary outcome to assess results if we had only included studies of higher methodological quality: analyzing only studies that (i) were of higher quality based on our prespecified threshold of quality assessment and that (ii) used multivariable logistic regression to control for potential confounders of ADR.

To estimate the relative risk difference, we calculated the relative risk (RR) from our pooled OR and 95% CIs using the equation provided by Zhang and Yu (25) where $RR = OR/((1-Po)+(Po\times OR))$, and Po is the pooled incidence of the outcome of interest (ADR or advanced ADR) among component studies in the reference group. From this relative risk and Po, we calculated the absolute risk differences as $(Po\times RR)-Po$. Publication bias was assessed by visual inspection of a funnel plot. No statistical test of funnel plot asymmetry was used because of the low power to detect a difference between chance and true asymmetry when <10 studies are included in a meta-analysis (24,26). RevMan 5.2 (The Cochrane Collaboration, Oxford, UK) was used for statistical analysis.

RESULTS

Search results

The initial search yielded 1,392 citations, of which a total of 910 unique citations remained after removal of duplicates (**Figure 1**). A recursive search of the reference sections of selected studies, practice guidelines, and pertinent review articles yielded an additional two articles. Of these 912 unique citations reviewed, 138 potentially relevant full articles were reviewed for inclusion. Six studies were excluded because they evaluated procedures other than colonoscopy, and 121 studies were excluded because they did not provide data allowing categorization of ADR or PDR into our prespecified bowel preparation strata, leaving 11 articles for inclusion.

Nine studies (27–35) met inclusion criteria for analysis of ADR for trichotomized comparisons (high/intermediate/low quality), although two studies (34,35) did not provide data allowing analysis of intermediate vs. low quality. Nine studies (27–33,36,37) met inclusion criteria for analysis of ADR for dichotomized comparisons (adequate vs. inadequate), four studies (27,28,31,34) met inclusion criteria for analysis of ADR for excellent vs. good-quality preparation, and four studies (28,30,31,35) met inclusion criteria for analyses of advanced ADR (for all comparisons). None of the 11 included studies provided data on detection of sessile serrated polyps/adenomas. **Table 1** summarizes the characteristics of the studies included and **Supplementary Figure S3** shows the scores for the six components of the quality assessment tool for the individual studies.

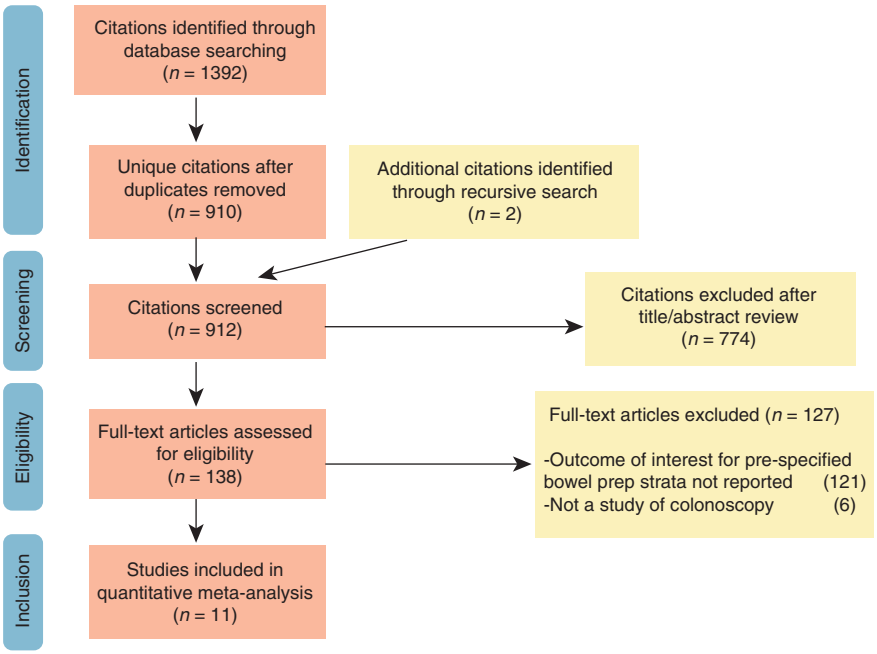

Figure 1. Flow diagram of study selection for systematic review and meta-analysis.

Primary end point: ADR in patients with intermediate-quality vs. high-quality preparation

Figure 2 displays the forest plot for the ADR of intermediate-quality vs. high-quality preparation for the nine individual studies (27–35), as well as the pooled analysis. Given significant heterogeneity, the pooled OR was calculated using a random-effects model. One study (32), a low-quality retrospective trial with 283 patients, reported a significant difference in favor of high-quality preparation (OR = 0.27 (0.18–0.41)), whereas another study (31), a high-quality prospective trial with 2,210 patients, reported a significant difference in favor of intermediate-quality preparation (OR = 1.21 (1.01–1.45)). Results of the pooled analysis showed no significant difference in ADR between intermediate-quality and high-quality bowel preparation (OR = 0.94 (0.80–1.10)). The estimated absolute risk difference in ADR was –1% (–3 to 2%; **Table 2**).

Prespecified sensitivity analyses of studies of higher methodologic quality based on our quality assessment tool (27,28,31,34) or based on the provision of results using multivariable logistic regression to control for confounding factors (29,33–35) revealed results similar to the primary analysis: OR = 0.98 (0.81–1.19) and OR = 1.04 (0.96–1.12), respectively. A *post hoc* sensitivity analysis in which we estimated ADR based on PDR using the lower bound of the 95% CI of the conversion factor (0.595) (22,23) showed no change in the results of the pooled analysis: OR = 0.94 (0.80–1.10). In our prespecified subgroup analyses, we found no significant difference in treatment effect between results from the seven studies providing ADR (27–29,31,32,34,35) and the two studies providing only PDR (30,33) ($P = 0.14$) or between the four studies including only screening/surveillance colonoscopies (27,28,31,34) and the five studies including colonoscopy for any indication (29,30,32,33,35) ($P = 0.81$). A funnel plot for our primary analysis is shown in **Figure 3**. Visual inspection shows no suggestion of publication bias to favor intermediate-quality preparation.

Secondary end points

On comparison of intermediate-quality vs. low-quality preparation (27–33) (**Figure 4**), ADR was significantly higher in the intermediate-quality group in four of the seven component studies (29–31,33), whereas comparison of high-quality vs. low-quality preparation (27–31,33–35) (**Figure 5**) reveals that ADR was significantly higher in the high-quality preparation group in four of the nine studies (29,30,34,35). Significant heterogeneity was also present in these analyses. ADRs were significantly higher with both intermediate-quality (OR = 1.39 (1.08–1.79)) and high-quality (OR = 1.41 (1.21–1.64)) preparation compared with low-quality preparation, with estimated absolute increases in ADR of 5% for both comparisons (**Table 2**).

Comparison of excellent- with good-quality preparation revealed no significant difference in ADR (OR = 1.04 (0.90–1.21); **Table 2**). Comparison of adequate (excellent, good, and fair) vs. inadequate (poor and insufficient) preparation revealed that ADR was significantly higher in the adequate group in five of the nine studies (29,30,32,33,37), as well as in the pooled analysis (OR = 1.30 (1.19–1.42); **Table 2**).

Advanced adenomas

Pooled analysis of the four studies (28,30,31,35) revealed no significant difference for comparison of intermediate-quality to high- or low-quality preparation, although a possible trend was seen for high- vs. low-quality preparation (OR = 1.21 (0.98–1.50); **Table 3**). Advanced ADR was significantly higher for adequate compared with inadequate preparation (OR = 1.30 (1.02–1.67)),

Table 1. Characteristics of studies included in meta-analysis

Study	Location	Study design	Study population	% Male	Mean age (s.d.)	Number of endoscopists	Number of colonoscopies	Cecal intubation rate	Overall ADR	Indications	Bowel Preparation Quality Scale	Prep quality prevalence	Quality assessment score[a]
Adler *et al.* (10)	Germany	Prospective	Screening colonoscopy	46.3%	64 (7)	21	11,186	98%	21.7%	Not given	Aronchick-based unvalidated 5-point ordinal scale	High: 88% Intermediate: 9% Low: 3%	6
Aslanian *et al.* (27)	New Haven, CT	Prospective	Screening colonoscopy	50.4%	58 (9)	7	502	100%	44.1%	Screening: 71% Symptom/ surveillance: 29%	Aronchick scale	High: 88% Intermediate: 9% Low: 3%	5
Bryant *et al.* (36)	Australia	Retrospective	Any colonoscopy	52.6%	Not given	Not given	1,785	92%	24.8%[b]	Screening/ surveillance: 34% Symptoms: 66%	Unvalidated ordinal scale	Adequate: 87% Inadequate: 13%	2
Chiu *et al.* (28)	China	Retrospective	Screening colonoscopy	52.7%	51 (11)	5	3,079	100%	16.3%	Screening: 100%	Aronchick scale	High: 53% Intermediate: 33% Low: 14%	5
De Jonge *et al.* (29)	The Netherlands	Retrospective/ prospective	Any colonoscopy	47.2%	59 (16)	117	4,800	92%	26.2%	Screening: 9% Surveillance: 15% Symptoms: 76%	Unvalidated ordinal scale	High: 81% Intermediate: 31% Low: 6%	4
Froehlich *et al.* (30)	Europe/ Canada	Prospective	Any colonoscopy	48.7%	58 (16)	208	5,832	88%	18.9%[b]	Screening: 10% Surveillance: 17% Symptoms: 73%	Aronchick-based unvalidated 5-point ordinal scale	High: 74% Intermediate: 16% Low: 10%	2
Gao *et al.* (38)	China	Prospective	Screening colonoscopy	68.9%	58 (10)	13	1,012	Not given	21.2%[b]	Screening: 100%	Boston Bowel Preparation Scale	Adequate: 87% Inadequate: 13%	3
Jover *et al.* (31)	Spain	Prospective	Screening colonoscopy	48.7%	59 (6)	Not given	4,539	95%	31.9%	Screening: 100%	Aronchick scale	High: 84% Intermediate: 14% Low: 2%	5
Perez *et al.* (32)	Peru	Retrospective	Any colonoscopy	33.6%	56 (15)	10	843	95%	19.9%	Screening: 52% Surveillance: 5% Symptoms: 43%	Unvalidated ordinal scale	High: 29% Intermediate: 45% Low: 26%	3
Radaelli *et al.* (33)	Italy	Prospective	Any colonoscopy	53.3%	60 (15)	673	12,835	81%	17.6%[b]	Screening: 14% Surveillance: 25% Symptoms: 61%	Unvalidated ordinal scale	High: 37% Intermediate: 47% Low: 16%	3
Sherer *et al.* (35)	Indianapolis, IN	Retrospective	Any colonoscopy	95.2%	63 (not given)	100	8,800	100%	Not given	Screening: 29% Surveillance: 27% Symptoms: 44%	Unvalidated ordinal scale	High: 59% Intermediate: 32% Low: 9%	3

ADR, adenoma detection rate; Prep, preparation.

[a]Adapted from Newcastle–Ottawa scale and QUADAS2 as described in Methods and Supplementary Figure S2. Supplementary Figure S3 shows the scores for the six components of this quality assessment tool for the individual studies.

[b]ADR estimated as polyp detection rate (PDR)×0.645.

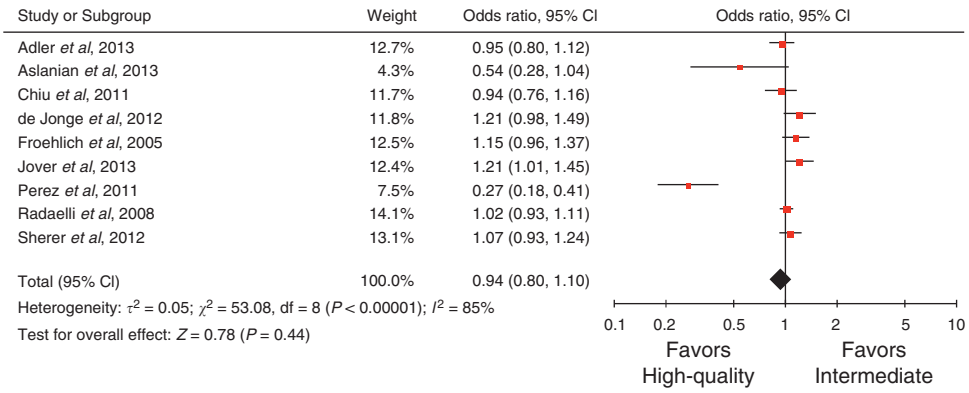

Figure 2. Comparison of adenoma detection rates for studies of intermediate-quality vs. high-quality bowel preparation. CI, confidence interval.

Table 2. Pooled odds ratio and estimated risk differences for comparisons of adenoma detection rates with different strata of bowel preparation quality

Comparison (number)	Adenoma detection rates			
	Number of studies	Odds ratio (95% CI)	Relative risk difference (95% CI)	Absolute risk difference (95% CI)
Intermediate quality (*N*=13,413) vs. high quality (*N*=34,211)	9[a]	0.94 (0.80, 1.10)	−5% (−17, 8%)	−1% (−3, 2%)
Intermediate quality (*N*=9,556) vs. low quality (*N*=3,699)	7[a]	1.39 (1.08, 1.79)	31% (7, 59%)	5% (1, 9%)
High quality (*N*=34,211) vs. low quality (*N*=4,899)	9[a]	1.41 (1.21, 1.64)	32% (17, 49%)	5% (3, 8%)
Adequate (*N*=31,047) vs. inadequate (*N*=4,058)	9[b]	1.30 (1.19, 1.42)	24% (15, 33%)	4% (2, 5%)
Excellent (*N*=6,794) vs. good (*N*=9,054)	4	1.04 (0.90, 1.21)	3% (−7, 14%)	1% (−2, 4%)

CI, confidence interval.
[a]Two studies reported only polyp detection rates.
[b]Four studies reported only polyp detection rates.

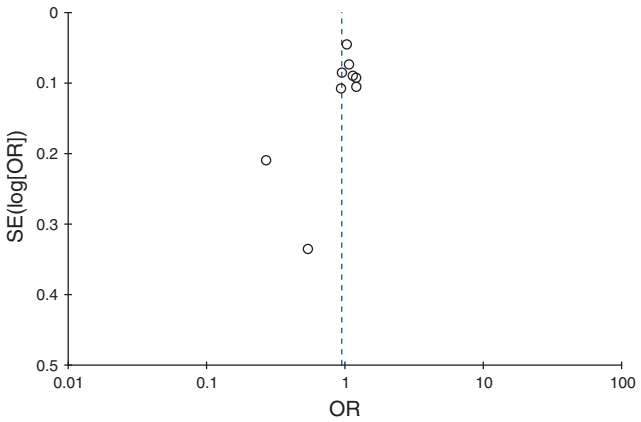

Figure 3. Funnel plot of studies in primary analysis of intermediate-quality vs. high-quality bowel preparation. OR, odds ratio.

with an estimated absolute increase in advanced ADR of 1% (0.1–3%; **Table 3**). Because one study's definition for advanced adenoma also included patients with ≥3 adenomas (31), we performed a *post hoc* sensitivity analysis without this study and found similar results: for advanced ADR with adequate vs. inadequate preparation, OR = 1.31 (1.00–1.71) and absolute difference = 2% (0–4%).

DISCUSSION

Documenting the quality of bowel preparation is recommended as a quality indicator for colonoscopy reporting (6,12). It seems obvious that better visualization of the colon would be associated with an improved ability to detect polyps, and numerous studies support this association, including our current meta-analysis. However, simply identifying the association that cleaner is better has limited clinical impact, as it does not answer the question of how clean is clean enough.

For screening or surveillance colonoscopy, the most important aspect of the bowel preparation quality is not the actual quality level, but whether the visualization was adequate to allow identification of polyps such that routine guideline-recommended surveillance intervals are appropriate (12). Although few would argue that

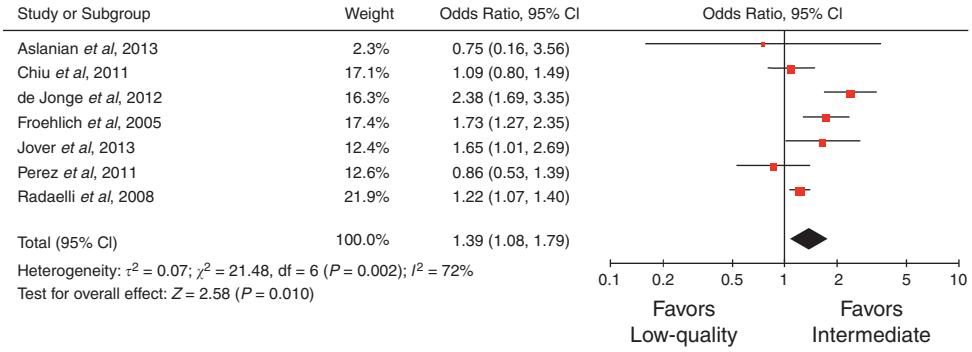

Figure 4. Comparison of adenoma detection rates for studies of intermediate-quality vs. low-quality bowel preparation. CI, confidence interval.

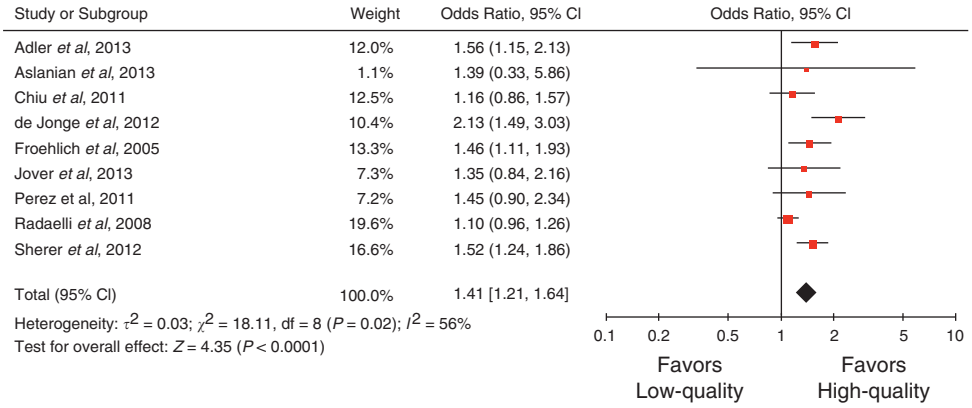

Figure 5. Comparison of adenoma detection rates for studies of high-quality vs. low-quality bowel preparation. CI, confidence interval.

Table 3. Pooled odds ratio and estimated risk differences for comparisons of advanced adenoma detection rates with different strata of bowel preparation quality

Comparison (number)	Advanced adenoma detection rates			
	Number of studies	Odds ratio (95% CI)	Relative risk difference (95% CI)	Absolute risk difference (95% CI)
Intermediate quality (*N*=8,517) vs. high quality (*N*=14,491)	4[a]	0.89 (0.69, 1.14)	−10% (−29, 13%)	−1% (−2, 1%)
Intermediate quality (*N*=8,517) vs. low quality (*N*=3,162)	4[a]	1.18 (0.70, 1.98)	17% (−29, 90%)	1% (−1, 4%)
High quality (*N*=14,491) vs. low quality (*N*=3,162)	4[a]	1.21 (0.98, 1.50)	20% (−2, 47%)	1% (−0.1, 2%)
Adequate (*N*=23,008) vs. inadequate (*N*=3,162)	4[a]	1.30 (1.02, 1.67)	28% (2, 62%)	1% (0.1, 3%)

CI, confidence interval.
[a]One study reported only polyp detection rates.

patients with low-quality bowel preparation should undergo repeat examination before determining appropriate surveillance intervals (11,12), the management of patients with intermediate-quality bowel preparation is uncertain. The results from our systematic review and meta-analysis demonstrate that ADR in patients with intermediate-quality bowel preparation is significantly higher than in patients with low-quality preparation, and it is not significantly different from ADR in patients with high-quality preparation. Our results confirm the need for early repeat colonoscopy for patients with low-quality bowel preparation, but suggest that patients with intermediate-quality preparation may be followed up at standard guideline-recommended surveillance intervals without the need for early repeat colonoscopy.

Closer assessment of studies comparing intermediate- with high-quality preparation reveals heterogeneous results with no significant difference in seven studies, a large prospective higher-quality

study showing significantly higher ADR with intermediate-quality preparation (31), and a small retrospective lower-quality study with a significantly higher ADR with high-quality preparation (32). The latter study reported that 30% of procedures had withdrawal times <6 min, potentially limiting its reliability. Although we cannot entirely exclude the possibility of a small decrement in ADR with intermediate as compared with high-quality preparation, our results are consistent with a nonsignificant 1% absolute decrease in ADR, with a maximum decrease of 3% based on the lower limit of the 95% CI. Furthermore, as compared with the result of our primary analysis of all studies (OR of 0.94), results of sensitivity analyses restricted to higher quality studies moved nonsignificantly closer to unity (ORs of 0.98 and 1.04)—in the direction of a smaller decrement or no decrement in ADR with intermediate- vs. high-quality bowel preparation.

The hypothesis that intermediate-quality preparation can be considered adequate for adenoma detection is also supported by other findings. ADR was significantly higher with both high-quality and intermediate-quality preparation compared with low-quality preparation, with similar ORs (1.41 and 1.39, respectively) and absolute increases (5% for both comparisons). These results indicate that 1 patient out of 20 with low-quality preparation will have no adenoma identified when adenomas would have been seen with either high- or intermediate-quality preparation.

Bowel preparation quality scoring methods commonly distinguish between excellent and good preparation. For example, the commonly used Aronchick scale has excellent and good categories, and the BBPS has analogous scores of 3 ("entire mucosa of colon segment seen well with no residual staining, small fragments of stool or opaque liquid") and 2 ("minor amount of residual staining, small fragments of stool and/or opaque liquid, but mucosa of colon segment seen well") for each colonic segment (20). Our results showed no significant increase in ADR with excellent vs. good preparation (OR of 1.04), suggesting that such a distinction may not be clinically necessary when assessing bowel preparation quality.

In evaluating advanced adenomas, the most clinically relevant lesions, our search yielded only four articles with analyzable data. Adequate-quality preparation was associated with significantly higher advanced ADR compared with inadequate preparation (OR = 1.30 (1.02–1.67)), whereas comparisons among the high-, intermediate-, and low-quality preparation categories did not yield significant differences. The lack of differences could have several explanations. First, advanced lesions are less prevalent, and thus we may not have had adequate power to identify differences in detection rates. Second, the definition of advanced lesions was heterogeneous among studies. Third, it could be hypothesized that larger lesions (≥1 cm) are less susceptible to being missed owing to the quality of preparation. For instance, after adjusting for age and gender among 93,004 complete colonoscopies in a US database, Harewood *et al.* (8) found "adequate" preparation to be associated with significantly higher detection of polyps ≤9 mm (OR 1.23 (1.19–1.28)) but no significant difference in detection of polyps >9 mm (OR 1.05 (0.98–1.11)). This study was excluded from our meta-analysis because it did not meet our predefined criteria for bowel preparation quality strata: it included "fair" and "fair with adequate examination" among the "inadequate" group.

In contrast, our data do suggest that bowel preparation quality, at least dichotomized as adequate or inadequate, does have a significant and potentially clinically relevant impact on the detection of advanced adenomas. Advanced adenomas are estimated to progress to cancer at a rate of 1% per year (38), with rates increasing over time. Using results from our study including a 4.2% prevalence of advanced adenoma and a 1–2% absolute increase in advanced ADR with adequate vs. inadequate preparation, we estimate that if repeat colonoscopy is not performed for 5 (if other adenomas are discovered) or 10 years (no adenomas discovered) instead of 3 years in the 1–2% of patients with missed advanced adenomas, this would result in an estimated 0.2–1.4 additional colorectal cancers per 1,000 patients undergoing screening or surveillance colonoscopy owing to inadequate bowel preparation. Assuming ~7 million patients in the United States undergoing colonoscopy for screening or surveillance annually (39), this would result in 1,400–9,800 interval cancers among these patients over the next 10 years resulting solely from missed lesions because of bowel preparation quality. These findings further support the recommendation for repeating the examination in patients with inadequate-quality preparation before determining an appropriate surveillance interval (12).

This study has limitations. There was heterogeneity in the populations evaluated and in the categorization of bowel preparation quality among component studies (**Table 1**). We attempted to minimize this heterogeneity by only including studies that allowed us to standardize bowel preparation quality definitions based on the five-tiered Aronchick classification (19). However, assessment of bowel preparation quality remains subjective, and even among "validated" bowel preparation scores some degree of interobserver variability remains (15,21). The Ottawa scale (21) and BBPS (15,20) have been developed as validated quantitative measures of bowel preparation quality. However, they have yet to be implemented universally and, as mentioned, only one of the studies we identified used one of these scoring systems. In addition, although it is most appropriate to score bowel preparation quality after washing to most accurately reflect mucosal visualization, we are unable to ensure that bowel preparation quality scores in the included studies all represent preparation quality after washing. If endoscopists recorded preparation quality as intermediate and then improved visualization with washing, the ADR for intermediate-quality preparation potentially might be increased, possibly lessening any difference in ADR between intermediate- and high-quality preparation. Identifying a simple method for defining bowel preparation quality after washing that can be universally understood and implemented will be critical to future studies evaluating bowel preparation quality and clinical outcomes.

Another source of heterogeneity included the reporting of polyp detection data in some studies and adenoma detection data in other studies. Francis *et al.* (23) and Boroff *et al.* (22) demonstrated that ADR can be accurately estimated by multiplying PDR by a conversion factor of 0.64 or 0.65, respectively, with high level of correlation. Thus, we felt that it was reasonable to include those studies

evaluating polyp-level data, multiplying raw polyp detection rates by 0.645 to estimate ADR. For studies that provided adjusted odds ratios, the use of PDR may be less of a concern because the conversion factor affects the numerator and denominator equally, cancelling itself out. In addition, we performed subgroup analysis including only studies using adenoma data and found no significant difference in our primary end point.

Finally, although ADR is the standard quality indicator currently used for colonoscopy, other measures, such as the total number of adenomas identified per colonoscopy, could potentially turn out to be a better quality metric and indicator of the development of interval cancers. For example, a less than optimal bowel preparation might allow the same number of patients to be identified with adenomas as compared with an excellent preparation but still lead to more missed adenomas per patient. The information provided in our component studies did not allow us to assess this end point. In addition, although ADR has been documented to be a surrogate for the development of interval colorectal cancer (7), studies assessing the effect of bowel preparation quality on the incidence of interval cancers would provide the most directly relevant information.

An intermediate-quality bowel preparation appears to be an important cause of early repeat colonoscopy despite the lack of information indicating a decreased ADR or increased rate of interval cancers. Ben-Horin *et al.* (13) reported that >60% of endoscopists recommended repeat colonoscopy at intervals earlier than 5 years in patients with normal colonoscopies, whereas Menees *et al.* (14) found that, of the 32% of average-risk patients with fair preparation and normal colonoscopy at initial screening, 57% received a recommendation for repeat colonoscopy within 5 years. Our data suggest, however, that if adequacy is measured by ADR, intermediate-quality preparation should be considered adequate for routine guideline-recommended surveillance intervals. In this era of value-driven health care, eliminating unnecessarily early screening/surveillance intervals in this large patient population could yield significant cost savings with minimal impact on quality, as measured by ADR. Further studies designed to address this issue should be conducted to corroborate this. Future research should focus on identifying a quantitative, reproducible measure that can identify the level of preparation quality that provides visualization adequate to have patients return at the standard, guideline-recommended surveillance intervals.

In summary, our systematic review and meta-analysis demonstrated no significant difference in ADR between high-quality and intermediate-quality bowel preparation, with similar significant increases in ADR for both high-quality and intermediate-quality preparation as compared with low-quality preparation. Furthermore, differentiating between excellent and good bowel preparation may not be clinically useful or necessary. Adequate preparation (including excellent, good, and fair preparation) is associated with a significantly higher ADR and advanced ADR as compared with inadequate preparation. Our results confirm the need for early repeat colonoscopy for low-quality bowel preparation, but suggest that patients with intermediate/fair preparation quality may be followed up at standard guideline-recommended surveillance intervals without significantly affecting quality as measured by ADR. Our results should be confirmed by future prospective studies using validated standardized bowel preparation scales.

CONFLICT OF INTEREST

Guarantor of the article: Loren Laine, MD.

Specific author contributions: B.T. Clark and T. Rustagi: study design, literature search, collection of data, analysis and interpretation of data, and drafting of the manuscript; L. Laine: study design, collection of data, analysis and interpretation of data, and drafting of the manuscript. All authors approved the final draft submitted.

Financial support: This study was supported by NIH T32 DK007017 (to B.T.C.)

Potential competing interests: None.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012;62:10–29.
2. Rex DK, Johnson DA, Anderson JC *et al.* American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. Am J Gastroenterol 2009;104:739–50.
3. Thiis-Evensen E, Hoff GS, Sauar J *et al.* Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. Scand J Gastroenterol 1999;34:414–20.
4. Winawer SJ, Zauber AG, O'Brien MJ *et al.* Randomized comparison of surveillance intervals after colonoscopic removal of newly diagnosed adenomatous polyps. The National Polyp Study Workgroup. N Engl J Med 1993;328:901–6.
5. Citarda F, Tomaselli G, Capocaccia R *et al.* Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. Gut 2001;48:812–5.
6. Rex DK, Petrini JL, Baron TH *et al.* Quality indicators for colonoscopy. Am J Gastroenterol 2006;101:873–85.
7. Kaminski MF, Regula J, Kraszewska E *et al.* Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010;362:1795–803.
8. Harewood GC, Sharma VK, de Garmo P. Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia. Gastrointest Endosc 2003;58:76–9.
9. Froehlich F, Wietlisbach V, Gonvers JJ *et al.* Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc 2005;61:378–84.
10. Adler A, Wegscheider K, Lieberman D *et al.* Factors determining the quality of screening colonoscopy: a prospective study on adenoma detection rates, from 12,134 examinations (Berlin colonoscopy project 3, BECOP-3). Gut 2013;62:236–41.
11. Lieberman D, Nadel M, Smith RA *et al.* Standardized colonoscopy reporting and data system: report of the Quality Assurance Task Group of the National Colorectal Cancer Roundtable. Gastrointest Endosc 2007;65:757–66.
12. Lieberman DA, Rex DK, Winawer SJ *et al.* Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2012;143:844–57.
13. Ben-Horin S, Bar-Meir S, Avidan B. The impact of colon cleanliness assessment on endoscopists' recommendations for follow-up colonoscopy. Am J Gastroenterol 2007;102:2680–5.
14. Menees SB, Kim HM, Elliott EE *et al.* The impact of fair colonoscopy preparation on colonoscopy use and adenoma miss rates in patients undergoing outpatient colonoscopy. Gastrointest Endosc 2013;78:510–6.
15. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc 2010;72:686–92.
16. Stroup DF, Berlin JA, Morton SC *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008–12.
17. Wells G SB, O'Connell J, Robertson J *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis [Website]. 2013 Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

18. Whiting PF, Rutjes AW, Westwood ME *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36.
19. Aronchick CA, Lipshutz WH, Wright SH *et al.* A novel tableted purgative for colonoscopic preparation: efficacy and safety comparisons with Colyte and Fleet Phospho-Soda. Gastrointest Endosc 2000;52:346–52.
20. Lai EJ, Calderwood AH, Doros G *et al.* The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. Gastrointest Endosc 2009;69:620–5.
21. Rostom A, Jolicoeur E. Validation of a new scale for the assessment of bowel preparation quality. Gastrointest Endosc 2004;59:482–6.
22. Boroff ES, Gurudu SR, Hentz JG *et al.* Polyp and adenoma detection rates in the proximal and distal colon. Am J Gastroenterol 2013;108:993–9.
23. Francis DL, Rodriguez-Correa DT, Buchner A *et al.* Application of a conversion factor to estimate the adenoma detection rate from the polyp detection rate. Gastrointest Endosc 2011;73:493–7.
24. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: www.cochrane-handbook.org.
25. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690–1.
26. Sterne JA, Sutton AJ, Ioannidis JP *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002.
27. Aslanian HR, Shieh FK, Chan FW *et al.* Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. Am J Gastroenterol 2013;108:166–72.
28. Chiu HM, Lin JT, Lee YC *et al.* Different bowel preparation schedule leads to different diagnostic yield of proximal and nonpolypoid colorectal neoplasm at screening colonoscopy in average-risk population. Dis Colon Rectum 2011;54:1570–7.
29. de Jonge V, Sint Nicolaas J, Cahen DL *et al.* Quality evaluation of colonoscopy reporting and colonoscopy performance in daily clinical practice. Gastrointest Endosc 2012;75:98–106.
30. Froehlich F, Wietlisbach V, Gonvers JJ *et al.* Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. Gastrointest Endosc 2005;61:378–84.
31. Jover R, Zapater P, Polania E *et al.* Modifiable endoscopic factors that influence the adenoma detection rate in colorectal cancer screening colonoscopies. Gastrointest Endosc 2013;77:381–9. e1.
32. Perez VP, Yamamoto JW, Nago AN *et al.* Quality indicators for colonoscopy in the Peruvian Japanese Policlinic: linked factors]. Acta Gastroenterol Latinoam 2011;41:288–95.
33. Radaelli F, Meucci G, Sgroi G *et al.* Technical performance of colonoscopy: the key role of sedation/analgesia and other quality indicators. Am J Gastroenterol 2008;103:1122–30.
34. Adler A, Wegscheider K, Lieberman D *et al.* Factors determining the quality of screening colonoscopy: a prospective study on adenoma detection rates, from 12,134 examinations (Berlin colonoscopy project 3, BECOP-3). Gut 2013;62:236–41.
35. Sherer EA, Imler TD, Imperiale TF. The effect of colonoscopy preparation quality on adenoma detection rates. Gastrointest Endosc 2012;75:545–53.
36. Bryant RV, Schoeman SN, Schoeman MN. Shorter preparation to procedure interval for colonoscopy improves quality of bowel cleansing. Intern Med J 2013;43:162–8.
37. Gao Y, Lin JS, Zhang HD *et al.* Pilot validation of the Boston bowel preparation scale in China. Dig Endosc 2013;25:167–73.
38. Stryker SJ, Wolff BG, Culp CE *et al.* Natural history of untreated colonic polyps. Gastroenterology 1987;93:1009–13.
39. Peery AF, Dellon ES, Lund J *et al.* Burden of gastrointestinal disease in the United States: 2012 update. Gastroenterology 2012;143:1179–87. e1–3.