

Improved detection of colorectal adenomas by high-quality colon cleansing



Authors

Cesare Hassan¹, Jonathan Manning², Marco Antonio Álvarez González³, Prateek Sharma⁴, Michael Epstein⁵, Raf Bisschops⁶

Institutions

- 1 Ospedale Nuovo Regina Margherita, Rome, Italy
- 2 Borders General Hospital, NHS Borders, Melrose, Berwickshire, UK
- 3 Hospital del Mar, Barcelona, Spain
- 4 Veterans Affairs Medical Center and University of Kansas School of Medicine, Kansas City, Kansas, United States
- 5 Investigative Clinical Research, Annapolis, Maryland, United States
- 6 KU Leuven, University Hospitals Leuven, Leuven, Belgium

submitted 13.12.2019

accepted after revision 9.4.2020

Bibliography

DOI <https://doi.org/10.1055/a-1167-1359> |
Endoscopy International Open 2020; 08: E928–E937
© Georg Thieme Verlag KG Stuttgart · New York
eISSN 2196-9736

Corresponding author

Cesare Hassan, Gastroenterology and Endoscopic Unit,
Nuovo Regina Margherita Hospital, Via Morosini 30, 00153
Roma RM, Italy
Fax: +447912248739
cesareh@hotmail.com

ABSTRACT

Background and study aims Reliable adenoma detection requires “adequate” bowel preparation. The adenoma detection rate (ADR) was assessed in patients with high-quality (stool-free) cleansing versus adequate cleansing.

Patients and methods This study was a post-hoc combined analysis of three randomized trials individually powered for cleansing quality assessment. Treatment-independent ADR was assessed versus colon cleansing quality by central readers using the Harefield Cleansing Scale (HCS) and the Boston Bowel Preparation Scale (BBPS). The number needed to treat (NNT) to find an additional patient with at least one adenoma was calculated for high-quality versus adequate-quality cleansing.

Results A total of 1749 patients were included. ADR increased with high-quality versus adequate-quality cleansing: HCS grade A versus B, 39% (94/242) versus 27% (336/1229); NNT=8.7; $P<0.001$. ADR also increased with high-quality versus uniform adequate segmental cleansing scores: HCS grade A versus uniform segmental scores 2, 39% (94/242) versus 26% (97/379); NNT=7.5; $P<0.001$. ADR increased with top-quality versus adequate segmental cleansing scores: HCS uniform segmental scores 4 versus 2, 54% (21/39) versus 26% (97/379); NNT=3.6; $P<0.001$. ADR increased with BBPS 9 versus 6, 43% (71/166) versus 26% (247/950); NNT=6.0; $P<0.001$. Right colon ADR increased with top-quality versus adequate cleansing: HCS 4 versus 2, 20% (25/122) versus 11% (121/1117); NNT=10.4; $P<0.001$ and BBPS 3 versus 2, 15% (42/284) versus 11% (130/1192); NNT=25.8; $P=0.033$.

Conclusions High-quality colon cleansing improves adenoma detection, and it should be a priority for bowel preparations for colonoscopy.

Introduction

Colorectal cancer (CRC) screening reduces the incidence and mortality of CRC [1–4]. Preventive efficacy of colonoscopy depends on complete visualization of colorectal mucosa [5–8].

The definition of a sufficient level of visualization remains unclear. Detection of colorectal neoplasia has been strictly

related to an adequate rate of bowel preparation. However, it is unclear what the most clinically relevant level of cleansing should be. The US Multi-Society Task Force on CRC defined by expert consensus adequate level bowel cleansing for colonoscopy as one that allows the detection of lesions >5 mm in size [9, 10]. However, an excellent level has been suggested to result in a higher detection [11]. The term “adequate” has only

recently been quantitatively defined [6]. On the Boston Bowel Preparation Scale (BBPS), a score of 2 per bowel segment was non-inferior to segmental cleansing scores of 3 in the detection of lesions >5 mm [6, 12]. Such a cleansing level still permits stool presence, albeit removable. A higher level of cleansing (overall BBPS score 7–9, or individual segmental score of 3; stool-free) was needed to improve the detection of sessile serrated adenomas and polyps [7]. While it has been suggested that detection of adenomas of all sizes, even diminutive (<5 mm) has the largest impact on advanced adenomas, and CRC incidence and mortality, no “adequate” level cleansing has been defined for reliable identification of high-risk patients (those with ≥ 3 adenomas) [9, 10, 13–16]. High-risk patients’ 10-year post-colonoscopy risk of advanced neoplasia is close to that for patients with advanced adenomas (17.7% versus 21.9%), and about three times the risk in patients with only 1–2 small adenomas (6.3%) [14].

To examine the superiority of high-quality versus adequate-quality colon cleansing for detection of neoplasia, we performed a post-hoc analysis of three very similarly designed clinical trials, where the level of cleansing was simultaneously assessed by two validated scales and two types of operator with different levels of experience of cleansing quality assessment. We then analyzed how the number of high-quality segments per patient is associated with the detection of colorectal neoplasia.

Patients and methods

Patients

A post-hoc analysis was performed using data collected from three clinical trials where the colon cleansing efficacy and safety of 1 L polyethylene glycol (PEG) NER1006 was assessed versus three standard bowel preparations. Each trial has been described in detail elsewhere [17–19]. All three trials were phase III, randomized, multicenter, colonoscopist- and central reader-blinded trials conducted in Europe and the US. Together, these trials are the largest prospective study program to date of colon cleansing quality.

The relationship between adenoma detection rate (ADR) and colon cleansing quality, particularly stool-free high-quality cleansing (Harefield Cleansing Scale [HCS] grade A or segmental score 3 or 4, or BBPS overall score 9 or segmental score 3) versus stool-containing adequate-quality cleansing (HCS grade B or segmental score 2, or BBPS overall score 6 (i.e. 2+2+2) or segmental score 2), was assessed. Patient-level data were combined, irrespective of treatment, provided patients had a readable colonoscopy video, recorded HCS and BBPS scores, and counts of adenomas for the overall and right colon (► Fig. 1). ADR and mean adenomas per patient (MAP) were analyzed versus the attained level of colon cleansing on two validated cleansing scales.

Patients were males and females aged 18 to 85 years who required a screening, surveillance, or diagnostic colonoscopy. NER1006 was assessed versus oral sulfate solution in the NOCT study, [17] 2 L PEG bowel preparation in the MORA study, [18] and sodium picosulfate/magnesium citrate in the DAYB study

[19]. The original alternative primary endpoints were overall cleansing success and right colon high-quality cleansing success using the HCS by central readers [20]. Site colonoscopists performed the colonoscopies and used the HCS (HCS site colonoscopist) while also recording polyp detection. Videos of the colonoscopies were then assessed by independent central readers who scored the cleansing quality using both the HCS (HCS central reader) and the BBPS (BBPS central reader) [12, 20]. Segmental scoring criteria on the HCS and BBPS are provided in ► Table 1. Adenomas were verified by pathology. ADR was a key secondary endpoint in the original trials. ADR was calculated as the percentage of patients in the population with at least one adenoma.

Study design

This post hoc analysis used patient-level data from all seven treatment arms of the DAYB, MORA and NOCT trials to analyze treatment effects on colonic segmental cleansing quality scores versus adenoma counts in the overall or right colon. Specifically, we focused on the attainment of high-quality versus adequate cleansing quality and on the differences between these two cleansing outcomes on the detection of adenomas. We performed these analyses on combined data from all three trials. Trial-specific data are also provided for completeness.

Assessments

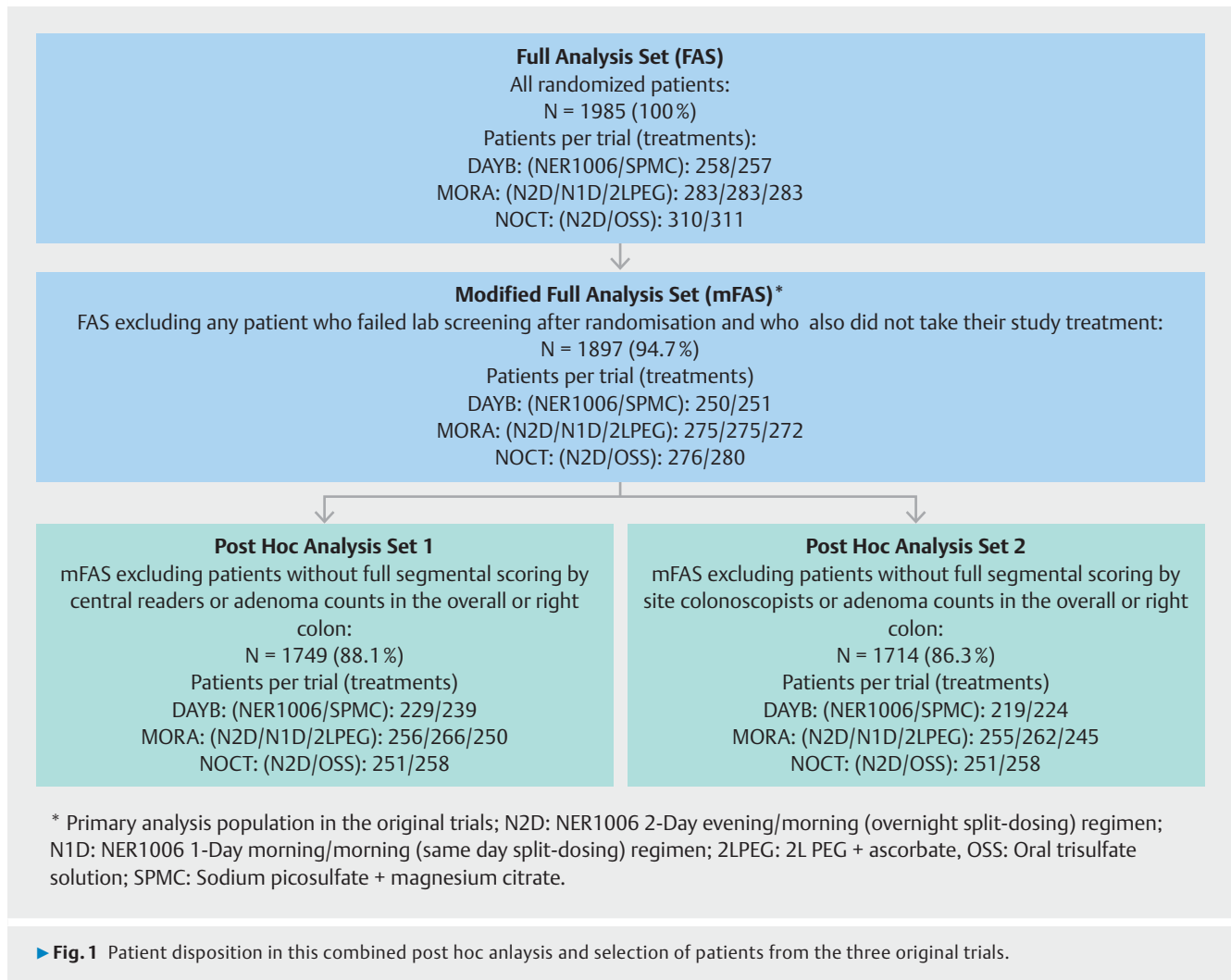
The primary aim of this post-hoc analysis was to determine the association between cleansing quality, assessed by treatment-blinded central readers using the HCS (Post hoc Analysis Set 1), and ADR in all patients, irrespective of bowel preparation received. The tested null hypothesis was: High-quality cleansing does not enable higher ADR than adequate level cleansing.

ADR was first compared in groups of patients with high-quality versus adequate-quality colon cleansing, as scored by HCS/central readers. High-quality cleansing was defined as stool-free (HCS segmental scores 3–4; grade A overall) where HCS 3 permits presence of ‘clear liquid’ while HCS score 4 is “empty and clean” (► Table 1). Adequate quality was defined as HCS score 2 which, like BBPS score 2, permits presence of removable stool. HCS cleansing quality data were stratified by:

1. Overall HCS cleansing grades.
2. Uniform segmental scores in all segments.
3. Right colon HCS segmental cleansing scores. The right colon represents the single-segment study level.

To illustrate the effect size with a clinically relevant measure, all ADR analyses of high-quality versus adequate-quality cleansing included a number needed to treat (NNT) calculation. The NNT to find another patient with at least one adenoma indicates how many patients are needed with stool-free rather than stool-containing cleansing to identify one more patient with an adenoma, either overall or in the right colon.

Complementary analyses were MAP according to each cleansing quality stratification, plus ADR and MAP versus each of BBPS central reader HCS (Post hoc Analysis Set 1) and HCS site colonoscopist HCS (Post hoc Analysis Set 2). On the BBPS, high-quality cleansing was defined as an overall score of 9 at



the whole colon level or score 3 at the segmental level. Adequate-quality cleansing was defined as BBPS 6 (2+2+2) and BBPS 2, respectively (► **Table 1**).

Statistics

A one-sided *t*-test was performed to assess the superiority of high-quality versus adequate-quality cleansing on adenoma detection. Because BBPS high-quality cleansing has already been shown to enable higher ADR than BBPS adequate quality cleansing, and all our post hoc analysis results on ADR and MAP also numerically favoured HQ over adequate, the one-sided *t*-test was considered most appropriate for assessment of superiority and it was thus used for the statistical comparisons.

The number needed to treat for an additional patient to benefit (NNT) was calculated by taking the reciprocal of the absolute risk reduction. In this analysis, the NNT was computed by taking the following steps: The difference between the adenoma detection rate (ADR) in the high-quality cleansing and the adequate-quality cleansing groups was calculated to obtain the absolute risk reduction between the two groups and the reciprocal of the absolute risk difference between the two groups was computed.

For the NNT to be clinically most useful, the NNT was only presented for results which also showed a statistical significance when using the two-sided *t*-test. Additionally, in order to quantify the two-sided uncertainty around the NNT, it was necessary to compute the two-sided 95% confidence interval (CI) for the NNT. This was obtained by taking the reciprocals of the values defining the confidence interval for the absolute risk reduction, in reverse order.

The two-sided *t*-test was also used to derive the 95% confidence interval for the MAP. For transparency, the two-sided 95% confidence interval is presented for all MAP differences.

All analyses were carried out using the statistical package R (v3.5.1), published by the R Foundation: <https://www.r-project.org/>.

Results

Patient characteristics

These analyses included 1749 patients with evaluable colonoscopy videos, recorded HCS and BBPS scores, and polyp and adenoma counts. Among them, 1714 had segmental scores by site colonoscopist (who, per study protocol, did not score all seg-

► **Table 1** Cleansing score criteria for the Harefield Cleansing Scale and the Boston Bowel Preparation Scale.

HCS[20]			BBPS[12]	
Required minimum score for overall grade	Score	Description	Score	Description
	0	Irremovable, heavy, hard stools	0	Unprepared colon segment with mucosa not seen due to solid stool that cannot be cleared
	1	Semi-solid, only partially removable stools	1	Portion of mucosa of the colon segment seen, but other areas of the colon segment not well seen due to staining, residual stool and/or opaque liquid
	2	Brown liquid/fully removable semi-solid stools	2	Minor amount of residual staining, small fragments of stool and/or opaque liquid, but mucosa of colon segment seen well
	3	Clear liquid	3	Entire mucosa of colon segment seen well with no residual staining, small fragments of stool or opaque liquid
	4	Empty and clean		

BBPS, Boston Bowel Preparation Scale; HCS, Harefield Cleansing Scale.

ments; with three exceptions, this happened only in patients with HCS grade D). Baseline characteristics for the 1749 patients by HCS central reader cleansing grade or total BBPS score are presented in ► **Table 2**. Trial-specific data on mean patient ages, polyp detection rate (PDR), adenoma detection rate (ADR), MPP and MAP in patients with adequate or high-quality cleansing are presented in ► **Supplementary Table 1**.

Adenoma detection versus central reader-assessed cleansing quality using the HCS is presented in ► **Table 3**. Adenoma detection versus central reader-assessed cleansing quality using the BBPS and site colonoscopist-assessed cleansing using the HCS are presented in ► **Table 4** and ► **Table 5**, respectively.

ADR versus HCS cleansing quality

ADR versus overall HCS cleansing grades

ADR was greater in patients with a high-quality cleansing grade versus an adequate only cleansing grade (► **Table 3**). HCS grade A achieved higher ADR than adequate-quality cleansing (grade B) (94/242, 39% versus 336/1229, 27%; $P < 0.001$). This finding indicates a positive effect on ADR when the cleansing grade improves, from adequate to high-quality cleansing.

ADR versus HCS grade A or uniform HCS segmental cleansing scores

ADR was higher with all segments scored 3 to 4 (HCS grade A) than with HCS 2 only (grade B without segmental scores HCS 3 to 4) (94/242, 39% versus 97/379, 26%; $P < 0.001$). ADR was also higher with uniform HCS 4 than with uniform HCS 2 (21/39, 54% versus 97/379, 26%; $P < 0.001$).

ADR versus right colon HCS segmental scores

ADR was numerically greater with right colon HCS scores 3 to 4 than with HCS 2 (56/410, 14% versus 121/1117, 11%; $P = 0.072$), and greater with HCS 4 than with HCS 2 (25/122, 20% versus 121/1117, 11%; $P = 0.001$).

NNT to find one more patient with at least one adenoma

The NNT was remarkably low for adenoma detection in the entire colon. To find one more patient with an adenoma in the overall colon, the results above suggest a need for only 3.6 to 8.7 patients with high-quality rather than adequate-quality cleansing.

At the single-segment level in the right colon, the NNT range was 10.4 to 35.4 patients.

MAP versus HCS cleansing quality

MAP largely reflected the findings for ADR (► **Table 3**). MAP increased numerically with HCS A versus HCS B (mean \pm standard deviation [SD]): 0.74 \pm 1.33 versus 0.58 \pm 1.67; $P = 0.080$).

ADR and MAP versus BBPS cleansing quality

ADR and MAP profiles were similar for BBPS and HCS. ADR improved with BBPS 9 versus 6 and uniform BBPS 3 versus 2 (identical analyses) (71/166, 43% versus 247/950, 26%; $P < 0.001$; ► **Table 4**). The right colon ADR improved with BBPS 3 versus 2: 42/284, 15% versus 130/1192, 11%; $P = 0.033$). NNTs of 6.0 patients were calculated for both overall and uniform segmental scores.

The right colon NNT was 25.8 patients, comparable (8.6) to the overall colon NNT if divided by the three BBPS segments and comparable to NNTs obtained with the HCS.

MAP improved with BBPS 9 versus 6 (0.86 \pm 1.41 versus 0.53 \pm 1.29; $P = 0.002$). Right colon MAP increased with BBPS 3 versus 2 (0.22 \pm 0.66 versus 0.16 \pm 0.54; $P = 0.047$).

ADR and MAP versus HCS cleansing quality by site colonoscopists

Site colonoscopists showed similar results to the central readers. ADR improved with HCS grade A versus B (297/946, 31% versus 147/561, 26%; $P = 0.016$; ► **Table 5**).

ADR improved numerically with only HCS scores 3 to 4 versus uniform HCS 2 (296/945, 31% versus 17/68, 25%; $P = 0.127$).

► Table 2 Baseline characteristics for the combined patient group from the three studies, by HCS grade and BBPS score group assessed by central readers (n = 1749).

Grade/score	Overall cleansing grade: Harefield Cleansing Scale				Overall cleansing score: Boston Bowel Preparation Scale				
	D	C	B	A	0–5	6 ¹	7 ²	8	9
Overall cleansing status	Failure		Success	High-quality	Failure	Success	Success	Success	High-quality
Stool-free colon	No	No	No	Yes	No	No	No	No	Yes
N	75	203	1229	242	338	950	167	128	166
Sex (male) n (%)	29 (38.7)	89 (43.8)	571 (46.5)	109 (45.0)	145 (42.9)	431 (45.4)	90 (53.9)	60 (46.9)	72 (43.4)
Age (y) mean ± SD	53.7 ± 15.2	54.9 ± 11.3	55.5 ± 12.1	54.7 ± 10.9	54.6 ± 12.5	55.0 ± 12.2	57.2 ± 11.0	56.8 ± 10.9	55.0 ± 11.0
Race n (%)									
White or Caucasian	74 (98.7)	197 (97.0)	1174 (95.5)	207 (85.5)	330 (97.6)	907 (95.5)	155 (92.8)	111 (86.7)	149 (89.8)
Other	1 (1.3)	6 (3.0)	55 (4.5)	35 (14.5)	8 (2.4)	43 (4.5)	7 (7.2)	17 (13.3)	17 (10.2)
BMI (kg/m ²) mean ± SD	27.3 ± 5.3	27.3 ± 5.3	27.7 ± 5.0	27.9 ± 5.1	27.4 ± 5.2	27.8 ± 5.1	27.7 ± 5.2	27.8 ± 5.0	27.8 ± 4.8
Reason for colonoscopy									
Screening	39 (52.0)	107 (52.7)	663 (53.9)	127 (52.5)	177 (52.4)	518 (54.5)	93 (55.7)	63 (49.2)	85 (51.2)
Surveillance	11 (14.7)	46 (22.7)	277 (22.5)	61 (25.2)	70 (20.7)	204 (21.5)	46 (27.5)	37 (28.9)	38 (22.9)
Diagnostic	25 (33.3)	50 (24.6)	289 (23.5)	54 (22.3)	91 (26.9)	228 (24.0)	28 (16.8)	28 (21.9)	43 (25.9)

BBPS, Boston Bowel Preparation Scale; BMI, body mass index; HCS, Harefield Cleansing Scale; SD, standard deviation

¹ BBPS score 2 + 2 + 2 (score cannot contain any segments scored 1).

² BBPS score 3 + 2 + 2 (score cannot contain any segments scored 1).

and with statistical significance with only HCS 4 versus 2 (47/117, 40% versus 17/68, 25%; $P=0.018$). Right colon ADR improved numerically with right colon HCS 3 to 4 or only 4 versus HCS 2 (127/1047, 12% and 41/278, 15% versus 56/508, 11%; $P=0.260$ and $P=0.065$, respectively). MAP increased with HCS grade A versus B (MAP ± SD: 0.68 ± 1.81 versus 0.47 ± 1.12 ; $P=0.007$). MAP was greater with HCS scores 3–4 versus uniform HCS 2 (0.68 ± 1.81 versus 0.44 ± 0.87 ; $P=0.027$) and with uniform HCS 4 vs 2 (1.20 ± 2.02 versus 0.44 ± 0.87 ; $P=0.002$).

The right colon MAP increased numerically with only HCS 3 to 4 versus 2 (0.17 ± 0.55 versus 0.15 ± 0.50 ; $P=0.223$) and with statistical significance with HCS 4 versus 2 (0.24 ± 0.69 versus 0.15 ± 0.50 ; $P=0.025$).

Discussion

The NER1006 phase III trials together form the largest prospective study program to date on bowel cleansing quality. Post-hoc analyses demonstrating superior high-quality cleansing with NER1006 suggest that it is now possible to routinely improve high-quality cleansing [21, 22]. We therefore examined the fundamental relationship between adenoma detection and increasing bowel cleansing quality in this combined post-hoc analysis.

High-quality cleansing, compared with an adequate level, is associated with an increased ADR. As this represents the main metric of quality for screening and surveillance colonoscopy,

we believe this finding is important. Results were reproducible when changing the scale for assessing colon cleansing and the experience of the operator, suggesting that the null hypothesis proposed when conducting these analyses can be rejected. Results were also confirmed when changing the study level from a patient to a segmental-based analysis and even when limiting our analysis to the right colon alone.

The results are clinically relevant. ADR is linked to future CRC disease risk and mortality rates [13, 23, 24] and our overall colon NNTs are in the single digit range. Single-segment NNTs are compatible with overall colon findings when adjusted for the number of segments in the entire colon. Patients with high-quality cleansing get better CRC prevention.

All analyses of ADR versus cleansing quality were repeated for MAP, which also increased with high-quality versus adequate cleansing in the overall colon as well as in the right colon. MAP improvement with HCS overall grade A versus B was numerically positive but it did not reach statistical significance. This was probably due to the added variability of how many, if any, adenomas each patient may have.

Our results support a recent cross-sectional study that showed an increased ADR with BBPS 3 versus 2 in over 4900 patients [25]. Our results are compatible with high-quality studies concluding lack of ADR improvement with higher-than-adequate bowel cleansing [6, 26]. Differences are likely to be attributable to small sample sizes or no standardized cleansing quality assessment [26].

► **Table 3** Adenoma detection versus HCS scoring by central readers.

	High-quality cleansing	Adequate-quality cleansing	High-quality versus adequate-quality: NNT [95% CI ¹] (ADR) or Difference [95% CI ¹] (MAP)	P (high-quality versus adequate-quality)
Adenoma detection rate in the overall colon, n/N (%)				
Overall HCS grade A versus B	94/242 (39%)	336/1229 (27%)	8.69 [5.63; 19.04]	<0.001
HCS grade A versus uniform HCS scores 2	94/242 (38.8%)	97/379 (26%)	7.55 [4.84; 17.08]	<0.001
Uniform HCS scores 4 versus 2	21/39 (54%)	97/379 (26%)	3.54 [2.33; 7.36]	<0.001
Adenoma detection rate in the right colon, n/N (%)				
Right colon HCS score 3–4 versus 2	56/410 (14%)	121/1117 (11%)	NA	0.072
Right colon HCS score 4 versus 2	25/122 (20%)	121/1117 (11%)	10.35 [6.38; 27.42]	0.001
Overall colon adenomas per patient, mean (SD)				
Overall HCS grade A versus B	0.74 (1.33)	0.58 (1.67)	0.16 [-0.06; 0.38]	0.080
HCS grade A versus uniform HCS 2	0.74 (1.33)	0.51 (1.18)	0.23 [0.03; 0.44]	0.013
Uniform HCS scores 4 versus 2	1.18 (1.80)	0.51 (1.18)	0.67 [0.26; 1.08]	0.014
Right colon adenomas per patient, mean (SD)				
Right colon HCS score 3–4 versus 2	0.20 (0.61)	0.16 (0.54)	0.04 [-0.02; 0.1]	0.113
Right colon HCS score 4 versus 2	0.34 (0.88)	0.16 (0.54)	0.18 [0.07; 0.29]	0.001
HCS, Harefield Cleansing Scale; MAP, mean adenomas per patient; NNT, number needed to treat; SD, standard deviation				
¹ Two-sided 95% confidence interval				

► **Table 4** Adenoma detection versus BBPS scoring by central readers.

	High-quality cleansing	Adequate-quality cleansing	High-quality versus adequate-quality: NNT [95% CI ¹] (ADR) or Difference [95% CI ¹] (MAP)	P (high-quality versus adequate-quality)
Adenoma detection rate in the overall colon, n/N (%)				
Overall BBPS score 9 versus 6	71/166 (43%)	247/950 (26%)	5.96 [4.14; 10.66]	<0.001
Uniform BBPS scores 3 versus 2	71/166 (43%)	247/950 (26%)	5.96 [4.14; 10.66]	<0.001
Adenoma detection rate in the right colon, n/N (%)				
Right colon BBPS score 3 versus 2	42/284 (15%)	130/1192 (11%)	NA	0.033
Overall colon adenomas per patient, Mean (SD)				
Overall BBPS score 9 versus 6	0.86 (1.41)	0.53 (1.29)	0.34 [0.12; 0.55]	0.002
Uniform BBPS scores 3 versus 2	0.86 (1.41)	0.53 (1.29)	0.34 [0.12; 0.55]	0.002
Right colon adenomas per patient, Mean (SD)				
Right colon BBPS score 3 versus 2	0.22 (0.66)	0.16 (0.54)	0.06 [-0.01; 0.14]	0.047
BBPS, Boston Bowel Preparation Scale; MAP, mean adenomas per patient; NNT, number needed to treat; SD, standard deviation				
¹ Two-sided 95% confidence interval				

The clinical potential of improved high-quality cleansing is compelling: adenomas and advanced adenomas are missed more frequently than previously believed, adenomas per positive index colonoscopy (i.e. multiple adenomas per patient) is the only factor independently associated with adenoma miss rates, and colonoscopies with low detection rates of multiple adenomas could cause 19,000 additional interval cancers in

the United States (US) alone [27]. Beyond the already mentioned 10-year risk being close to that of baseline advanced adenomas, a recent registry study in over 40,000 patients showed that the only risk factor with an odds ratio of >2 for an advanced adenoma detection during the first surveillance was adenoma multiplicity at baseline [14].

► **Table 5** Adenoma detection versus HCS scoring by site colonoscopists.

	High-quality cleansing	Adequate-quality cleansing	High-quality versus adequate-quality: NNT [95% CI ¹] (ADR) or Difference [95% CI ¹] (MAP)	P (high-quality versus adequate-quality)
Adenoma detection rate in the overall colon, n/N (%)				
Overall HCS grade A versus B	297/946 (31%)	147/561 (26%)	19.26 [10.05; 231.96]	0.016
HCS grade A versus uniform HCS scores 2	296/945 (31%)	17/68 (25%)	NA	0.127
Uniform HCS scores 4 versus 2	47/117 (40%)	17/68 (25%)	6.59 [3.4; 104.96]	0.018
Adenoma detection rate in the right colon, n/N (%)				
Right colon HCS score 3–4 versus 2	127/1047 (12%)	56/508 (11%)	NA	0.260
Right colon HCS score 4 versus 2	41/278 (15%)	56/508 (11%)	NA	0.065
Overall colon adenomas per patient, mean (SD)				
Overall HCS grade A versus B	0.68 (1.81)	0.47 (1.12)	0.21 [0.04; 0.37]	0.007
HCS grade A versus uniform HCS scores 2	0.68 (1.8)	0.44 (0.87)	0.24 [-0.2; 0.67]	0.027
Uniform HCS scores 4 versus 2	1.20 (2.02)	0.44 (0.87)	0.76 [0.25; 1.27]	0.002
Right colon adenomas per patient, mean (SD)				
Right colon HCS score 3–4 versus 2	0.17 (0.55)	0.15 (0.50)	0.02 [-0.04; 0.08]	0.223
Right colon HCS score 4 versus 2	0.24 (0.69)	0.15 (0.50)	0.08 [0; 0.17]	0.025
HCS, Harefield Cleansing Scale; MAP, mean adenomas per patient; NNT, number needed to treat; SD, standard deviation				
¹ Two-sided 95% confidence interval				

The role of the assessor performing segmental scoring seems to affect the quantitative relationship between ADR and colon cleansing quality on the HCS. Quantitatively, the NNT was higher for cleansing quality assessed by site colonoscopist versus central reader (15.8 to 19.3 versus 7.5 to 8.7). These results are consistent with previously published findings that site colonoscopists allocate higher segmental scores than central readers [26]. Endoscopy practitioners, however, will still experience ADR improvement with higher cleansing quality.

Results were highly reproducible between HCS and BBPS. Choosing either of these two validated scales has no impact on the positive relationship between ADR and colon cleansing quality. Both the HCS and the BBPS may evaluate accurately high-quality bowel preparation and they both have their respective strengths and limitations. While scales for assessment of bowel preparation quality for CRC screening colonoscopy have improved, establishing a standard, validated scale is essential to optimise CRC colonoscopy screening. BBPS is, however, relatively easy to use in clinical practice and with more publications versus the HCS on PubMed by an order of magnitude, the BBPS is clearly more widely used. BBPS therefore emerges as the best available option and it should be recommended as the current standard for use in clinical practice. Limitations to these analyses include their post-hoc nature and the inclusion of multiple indications for colonoscopy. Having fewer patients available for the HCS uniform analysis compared with the BBPS was also a limitation in the power of these analyses. This was due to the nature of the scales, where patients were required to have five uniform segmental scores to be included

in the HCS uniform analysis, as opposed three for the BBPS. Lesions were not characterized, so there is no available information on their size or morphology.

The strengths of these analyses include the dosing regimen- and treatment-independent analysis of a large number of patients and the very similar strict randomized trial setting that was used across all three phase III trials across multiple centres and countries to generate the data. The fact that one of the trials, DAYB, examined a less effective dosing regimen (day before dosing), reduced the overall success rates in this analysis but it also provided valuable patient level data in patients with only floor-level adequate cleansing. Use of treatment-blinded site colonoscopist and central readers is a more rigorous assessment method than is used in everyday clinical practice, and it probably reduced the influence of subjective bias as part of bowel cleansing assessment. Also, the ability to examine the relationship between cleansing grade and lesion detection was strengthened through the use of two different validated scales for bowel cleansing, each with different assessment methods [12, 20]. Our analyses provided complementary information. ADR indicates the number of patients who could benefit from improved cleansing, while the MAP indicates potential benefits for each patient with detected adenomas.

Conclusions

ADR improves with stool-free (high-quality) versus stool-containing (adequate-quality) cleansing scores on both the HCS and the BBPS, whether assessed by local or central experts.

High-quality cleansing improved the detection of multiple adenomas per patient, a measure of increasing importance for CRC prevention. We may have identified a new clinical risk that reliance on adequate-quality cleansing could lead to missed adenomas. Large randomized clinical trials hypothesizing improved ADR and MAP with stool-free cleansing must establish the quantitative benefits. Until then, high-quality cleansing is encouraged for pre-colonoscopy bowel preparation. Attention to bowel preparation is imperative and maximal cleansing efficacy should be the clinical priority.

Acknowledgements

The authors thank all the investigators, trial personnel, and participating patients for their contributions to the three underlying clinical trials that enabled our post hoc analysis. They also specifically thank those authors of the primary trial publications who did not also author the current manuscript, for their pivotal scientific contributions: The MORA trial publication authors Lucy B. Clayton and Richard Ng Kwet Shing, the NOCT trial publication authors Michael P. DeMicco, Lucy B. Clayton, and Jeff Pilot; and the DAYB trial publication authors Stefan Schreiber, Daniel C. Baumgart, Joost P.H. Drenth, Rafal S. Filip, Lucy B. Clayton, Kerry Hylands, and Alessandro Repici. Consultancy was provided by Dr. Juha Halonen and Lucy Clayton at Norgine Ltd, who also funded the writing support by Rory Elsome.

Competing interests

Dr. Hassan was an investigator in the DAYB study and has received honoraria from Norgine Ltd. for advisory board attendance. Dr. Manning was an investigator in the MORA study and has received honoraria from Norgine Ltd. for advisory board attendance and clinical conference attendance as a presenting author. Dr. Álvarez-González was an investigator in the MORA study and has received honoraria from Norgine Ltd. for advisory board attendance and from Casen-Recordati for speaking and teaching. Dr. Sharma has served as a consultant for Boston Scientific and received grants from CDx labs, US Endoscopy, and Medtronic. Dr. Epstein was an investigator in the NOCT study and has acted as a safety advisor for Aspire Bariatrics, a consultant for Zx Pharma and IMHealthScience, and a speaker for Daiichi Sankyo and Pfizer. Dr. Bisschops was an investigator in the MORA study and has received honoraria from Norgine Ltd. for speaking and advisory board attendance. This study was funded by Norgine Ltd. (Harefield, UK).

References

- [1] Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; 348: g2467
- [2] Lin JS, Piper MA, Perdue LA et al. Screening for Colorectal Cancer: A Systematic Review for the U.S. Preventive Services Task Force. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016: Available from Internet (cited 2017 Apr 25): <http://www.ncbi.nlm.nih.gov/books/NBK373584/>
- [3] Doubeni CA, Corley DA, Quinn VP et al. Effectiveness of screening colonoscopy in reducing the risk of death from right and left colon cancer: a large community-based study. *Gut* 2018; 67: 291–298
- [4] Nishihara R, Wu K, Lochhead P et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med* 2013; 369: 1095–1105
- [5] Oh CH, Lee CK, Kim J-W et al. Suboptimal bowel preparation significantly impairs colonoscopic detection of non-polypoid colorectal neoplasms. *Dig Dis Sci* 2015; 60: 2294–2303
- [6] Clark BT, Protiva P, Nagar A et al. Quantification of Adequate bowel preparation for screening or surveillance colonoscopy in men. *Gastroenterology* 2016; 150: 396–405; quiz e14–15
- [7] Clark BT, Laine L. High-quality bowel preparation is required for detection of sessile serrated polyps. *Clin Gastroenterol Hepatol* 2016; 14: 1155–1162
- [8] Anderson JC, Butterly LF, Robinson CM et al. Impact of fair bowel preparation quality on adenoma and serrated polyp detection: data from the New Hampshire colonoscopy registry by using a standardized preparation-quality rating. *Gastrointest Endosc* 2014; 80: 463–470
- [9] Rex DK, Bond JH, Winawer S et al. Quality in the technical performance of colonoscopy and the continuous quality improvement process for colonoscopy: recommendations of the U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2002; 97: 1296–1308
- [10] Lieberman DA, Rex DK, Winawer S 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: 844–857
- [11] Pontone S, Hassan C, Maselli R et al. Multiple, zonal and multi-zone adenoma detection rates according to quality of cleansing during colonoscopy. *United Eur Gastroenterol J* 2016; 4: 778–783
- [12] 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: 620–625
- [13] Meester RGS, Doubeni CA, Lansdorp-Vogelaar I et al. Variation in adenoma detection rate and the lifetime benefits and cost of colorectal cancer screening: a microsimulation model. *JAMA* 2015; 313: 2349
- [14] Lieberman D, Sullivan BA, Hauser ER et al. Baseline colonoscopy findings associated with 10-year outcomes in a screening cohort undergoing colonoscopy surveillance. *Gastroenterology*; 2019: Available from Internet (cited 2019 Aug 13): <https://linkinghub.elsevier.com/retrieve/pii/S0016508519411499>
- [15] Bonnington S, Sharp L, Rutter M. Post-polypectomy surveillance in the English bowel cancer screening programme: results of first surveillance. Prague, Czech Republic: *Endoscopy*; 2019: S1–S273. Available from Internet (cited 2019 Aug 13): <http://www.thieme-connect.de/DOI/DOI?10.1055/s-0039-1681511>
- [16] Bonnington S, Sharp L, Rutter M. Post-polypectomy surveillance in the English bowel cancer screening programme: multivariate logistic regression of factors influencing advanced adenoma detection at first surveillance. Prague, Czech Republic: *Endoscopy*; 2019: S1–S273. Available from Internet (cited 2019 Aug 13): <http://www.thieme-connect.de/DOI/DOI?10.1055/s-0039-1681622>
- [17] DeMicco MP, Clayton LB, Pilot J et al. Novel 1 L polyethylene glycol-based bowel preparation NER1006 for overall and right-sided colon cleansing: a randomized controlled phase 3 trial versus trisulfate. *Gastrointest. Endosc* 2018; 87: 677–687.e3
- [18] Bisschops R, Manning J, Clayton LB et al. Colon cleansing efficacy and safety with 1 L NER1006 versus 2 L polyethylene glycol + ascorbate: a randomized phase 3 trial. *Endoscopy* 2019; 51: 60–72

- [19] Schreiber S, Baumgart DC, Drenth JPH et al. Colon cleansing efficacy and safety with 1 L NER1006 versus sodium picosulfate with magnesium citrate: a randomized phase 3 trial. *Endoscopy* 2019; 51: 73–84
- [20] Halphen M, Heresbach D, Gruss H-J et al. Validation of the Harefield Cleansing Scale: a tool for the evaluation of bowel cleansing quality in both research and clinical practice. *Gastrointest. Endosc* 2013; 78: 121–131
- [21] Repici A, Amlani B, Uebel P et al. P0159 Superior high-quality colon cleansing with 1L NER1006 versus sodium picosulfate + magnesium citrate, 2L polyethylene glycol + ascorbate, or oral sulfate solution: Post hoc combined analysis of three randomised phase 3 clinical trials. *United Eur Gastroenterol J* 2018; 6: A183
- [22] Repici A, Coron E, Sharma P et al. Improved high-quality colon cleansing with 1L NER1006 versus 2L polyethylene glycol + ascorbate or oral sulfate solution. *Dig Liver Dis* 2019; 51: 1671–1677
- [23] Corley DA, Jensen CD, Marks AR et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; 370: 1298–1306
- [24] Kaminski MF, Wieszczy P, Rupinski M et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. *Gastroenterology* 2017; 153: 98–105
- [25] Guo R, Wang Y-J, Liu M et al. The effect of quality of segmental bowel preparation on adenoma detection rate. *BMC Gastroenterol*; 2019; 19: Available from Internet (cited 2019 Aug 13): <https://bmkgastroenterol.biomedcentral.com/articles/10.1186/s12876-019-1019-8>
- [26] Ell C, Fischbach W, Bronisch H-J et al. Randomized trial of low-volume PEG solution versus standard PEG + electrolytes for bowel cleansing before colonoscopy. *Am J Gastroenterol* 2008; 103: 883–893
- [27] Zhao S, Wang S, Pan P et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. *Gastroenterology* 2019; 156: 1661–1674.e11

► **Supplementary Table 1** Mean patient age, PDR, ADR, MPP and MAP in patients who attained HCS grades A (high-quality) or B (successful) in the combined patient group assessed by central readers (n = 1749), and by individual phase 3 trial DAYB, MORA and NOCT.

HCS Grade	Combined phase 3		DAYB		MORA		NOCT	
	HCS B	HCS A	HCS B	HCS A	HCS B	HCS A	HCS B	HCS A
Patients, n	1229	242	279	5	582	138	368	99
Age (years), mean	55.5	54.7	53.7	57.4	55.1	54.5	57.5	54.9
PDR, %	43.6%	50.8%	37.6%	20.0%	43.3%	51.4%	48.6%	51.5%
ADR, %	27.3%	38.8%	20.1%	0.0%	26.1%	37.0%	34.8%	43.4%
MPP, mean	1.12	1.35	0.90	0.20	1.18	1.54	1.20	1.15
MAP, mean	0.58	0.74	0.37	0.00	0.54	0.69	0.82	0.86

HCS, Harefield Cleansing Scale; PDR, polyp detection rate; ADR, adenoma detection rate; MPP, mean polyps per patient; MAP, mean adenomas per patient