Clean Colon Software Program (CCSP), Proposal of a standardized Method to quantify Colon Cleansing During Colonoscopy: Preliminary Results

(Article begins on next page)

Clean Colon Software Program (CCSP), Proposal of a standardized Method to quantify Colon Cleansing During Colonoscopy: Preliminary Results

Authors   Erik Rosa-Rizzotto[1], Adrian Dupuis[2], Ennio Guido[1], Diego Caroli[3], Fabio Monica[4], Daniele Canova[4], Erica Cervellin[5], Renato Marin[5], Cristina Trovato[6], Cristiano Crosta[6], Silvia Cocchio[7], Vincenzo Baldo[7], Franca De Lazzari[1]

Institutions   Institutions are listed at the end of article.

Corresponding author
Dr Erik Rosa-Rizzotto
Gastroenterology Unit, St
Anthony Hospital
Via Facciolati 73, 35127
Padova
Italy
Phone: ++39 049 821 7523
Fax: ++39 049 821 6741
erik.rosarizzotto@sanita.
padova.it

Background and study aims: Neoplastic lesions can be missed during colonoscopy, especially when cleansing is inadequate. Bowel preparation scales have significant limitations and no objective and standardized method currently exists to establish colon cleanliness during colonoscopy. The aims of our study are to create a software algorithm that is able to analyze bowel cleansing during colonoscopies and to compare it to a validate bowel preparation scale.

Patients and methods: A software application (the Clean Colon Software Program, CCSP) was developed. Fifty colonoscopies were carried out and video-recorded. Each video was divided into 3 segments: cecum-hepatic flexure (1st Segment), hepatic flexure-descending colon (2nd Segment) and rectosigmoid segment (3rd Segment). Each segment was recorded twice, both before and after careful cleansing of the intestinal wall. A score from 0 (dirty) to 3 (clean) was then assigned by CCSP. All the videos were also viewed by four endoscopists and colon cleansing was established using the Boston Bowel Preparation Scale. Interclass correlation coefficient was then calculated between the endoscopists and the software.

Results: The cleansing score of the prelavage colonoscopies was 1.56 ± 0.52 and the postlavage one was $2,08 \pm 0,59$ (P < 0.001) showing an approximate 33.3 % improvement in cleansing after lavage. Right colon segment prelavage (0.99 ± 0.69) was dirtier than left colon segment prelavage (2.07 ± 0.71). The overall interobserver agreement between the average cleansing score for the 4 endoscopists and the software pre-cleansing was 0.87 (95 % CI, 0.84 – 0.90) and post-cleansing was 0.86 (95 % CI, 0.83 – 0.89).

Conclusions: The software is able to discriminate clean from non-clean colon tracts with high significance and is comparable to endoscopist evaluation.

Introduction
▼

Colonoscopy is the diagnostic point of reference for colorectal cancer screening programs. Good bowel preparation is essential to produce a good-quality colonoscopy and it allows the detection of preneoplastic colon lesions [1 – 4].

Assessment of bowel preparation is entrusted to the endoscopist who is carrying out the endoscopy. Thus it is thus clearly subjective and associated with an infinite number of variables. Most bowel preparation scales have not been shown to be valid or reliable. Moreover, the few bowel preparation scales that have been validated have significant limitations, including an inability to distinguish among bowel preparations that adequately cleanse a high percentage of colons [5 – 9]. In the last position paper of the European Society of Gastrointestinal Endoscopy (ESGE), Hassan et al. drew up evidence- and consensus-based guidelines on bowel preparation for colonoscopy [10]. The main problem during the review of all randomized controlled trials (RCTs) and meta-analyses was that none of the studies were comparable in regard to scales for bowel cleansing. This is because no valid, reliable, internationally-approved rating scale exists for evaluation of the quality of bowel cleansing. Some scales – such as the Aronchic Scale [5], the Ottawa Scale [6], the Boston Bowel-Preparation Scale [7], the Harefield Cleansing Scale [8], and the Chicago scale [9] – have been proposed and validated, but they are difficult to apply in clinical practice and operator-dependent.

A computer-assisted clean-colon scale would be an attractive alternative to the current proposed scales because it could potentially be a reliable and more objective rating scale un affected by the endoscopist. The aims of our study are to create a software algorithm that is able to analyze
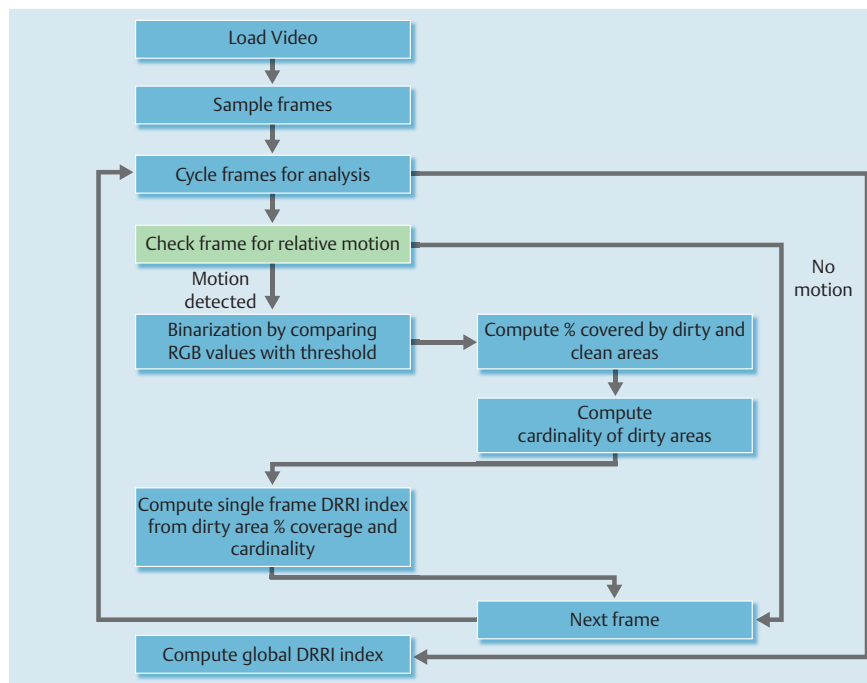
Fig. 1   Flowchart of the software program

bowel cleansing during colonoscopies and to correlate it with a scale for colon cleansing that has been validated.

Materials and methods
▼

Software application
The Clean Colon Software Program (CCSP) is a mathematical algorithm applicable to all Personal Computers that objectively analyzes bowel cleansing during colonoscopy examinations and results in a numerical scoring (Dupuis-Rosa-Rizzotto Index – DRRI) indicative of the degree of bowel cleanliness.

Video-recordings of colonoscopies were analyzed using a threshold analysis by a software program that first sampled a number of frames (depending on a frequency parameter related to the length of the video) and then each frame wa assessed to distinguish clean from dirty zones, as explained in ◉ **Fig. 1**. "Clean pixels" (e.g., red, pink, violet) were separated from the "dirty" ones (e.g., yellow, brown or green).

The threshold was that value which, compared with a precise calculation based on normalized Red-Green-Blue (RGB) values, made possible a distinction between the two types of zones: clean and dirty (◉ **Fig. 2**). RGB coloring is predicated on the physical principle that any spectral color can be generated by mixing the three basic colors. For example, if red and green are mixed at full intensity, bright yellow is generated, whereas a mixture of red and blue results in violet. If the three basic colors are mixed at different intensities, the whole spectrum of rainbow colors can be created.

The sum of pixel values of the two zones contributed to the calculation of the cleanliness index for each frame being analyzed (with values running from 0 to 3, and 3 representing the highest degree of cleanliness). The average of the indices for all of the frames resulted in an overall cleanliness index for a particular video (◉ **Fig. 3**).

Dupuis-Rosa-Rizzotto Index – DRRI
DRRI is calculated as follows:

Each frame is binarized, pixel by pixel:

$P = (0,0,0)$ if $(PR – PG) / (PG-PB) * 100 > TH$ (dirty areas)

$P = (255,255,255)$ if $(PR – PG) / (PG-PB) * 100 ≤ TH$ (clean areas)

Where PR, PG, PB are respectively the component of red, yellow, and blue color value of the pixel, and TH is the threshold value empirically determined by the characteristics of the device (for this study set to TH = 400)

% of the area with dirty pixels (Acovered) and the number of distinct areolas present in the frame (Adistinct) are calculated from a binarized frame.

If N sampled frame, with k from 1 to N, was estimated, the index of the κ-th frame resulted:

DRRIκ : 3 if $max(Acovered – 2* Adistinct, 0) <= TH1$

2 if $TH1 < max(Acovered – 2* Adistinct, 0) ≤ TH2$

1 if $TH2 < max(Acovered – 2* Adistinct, 0) ≤ TH3$

0 if $max(Acovered – 2* Adistinct, 0) > TH3$

TH1, TH2 e TH3 are threshold values that can be set. For this study, they were set to TH1 = 0, TH2 = 3, TH3 = 20

The final value of DRRI is the average of DRRIκ sampled frames:

$$DRRI = \frac{\sum_{\kappa=1}^{N} DRRI\kappa}{N}$$

Colonoscopy video collection
Fifty colonoscopy videos were collected in five different Italian Endoscopy Units (Padova, Milano, Bassano del Grappa, Dolo, Chioggia). Every video had a minimum resolution of at least 720 × 576 pixels and was saved in a digital video format (.avi file). The videos are available at www.youtube.com/user/PadovaCCSP. All of the videos were anonymized and patients gave consent for recording of their colonoscopies.

Colonoscopy videos were randomly collected by seven skilled endoscopists, using high-definition endoscopes, during routine and colorectal cancer screening sessions. Operative colonosco-

Fig. 2 Colon Cleansing Level. At left is the original frame and at right is the elaborated image with white pixel "dirty" areas and DRRI corresponding index.
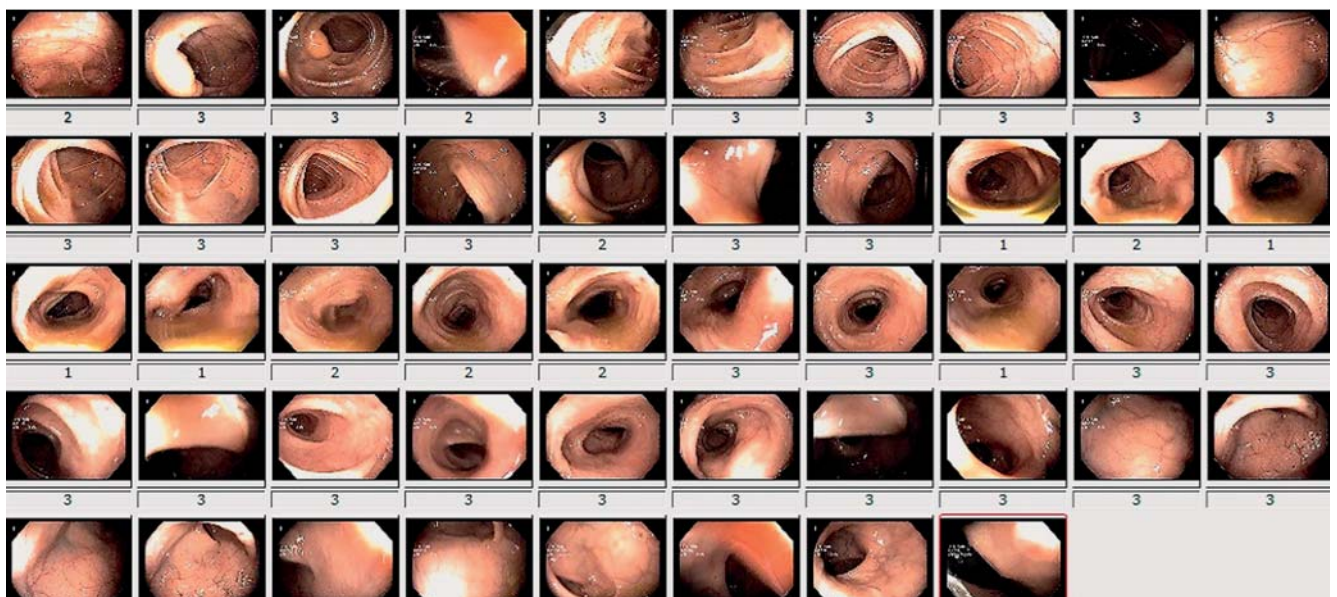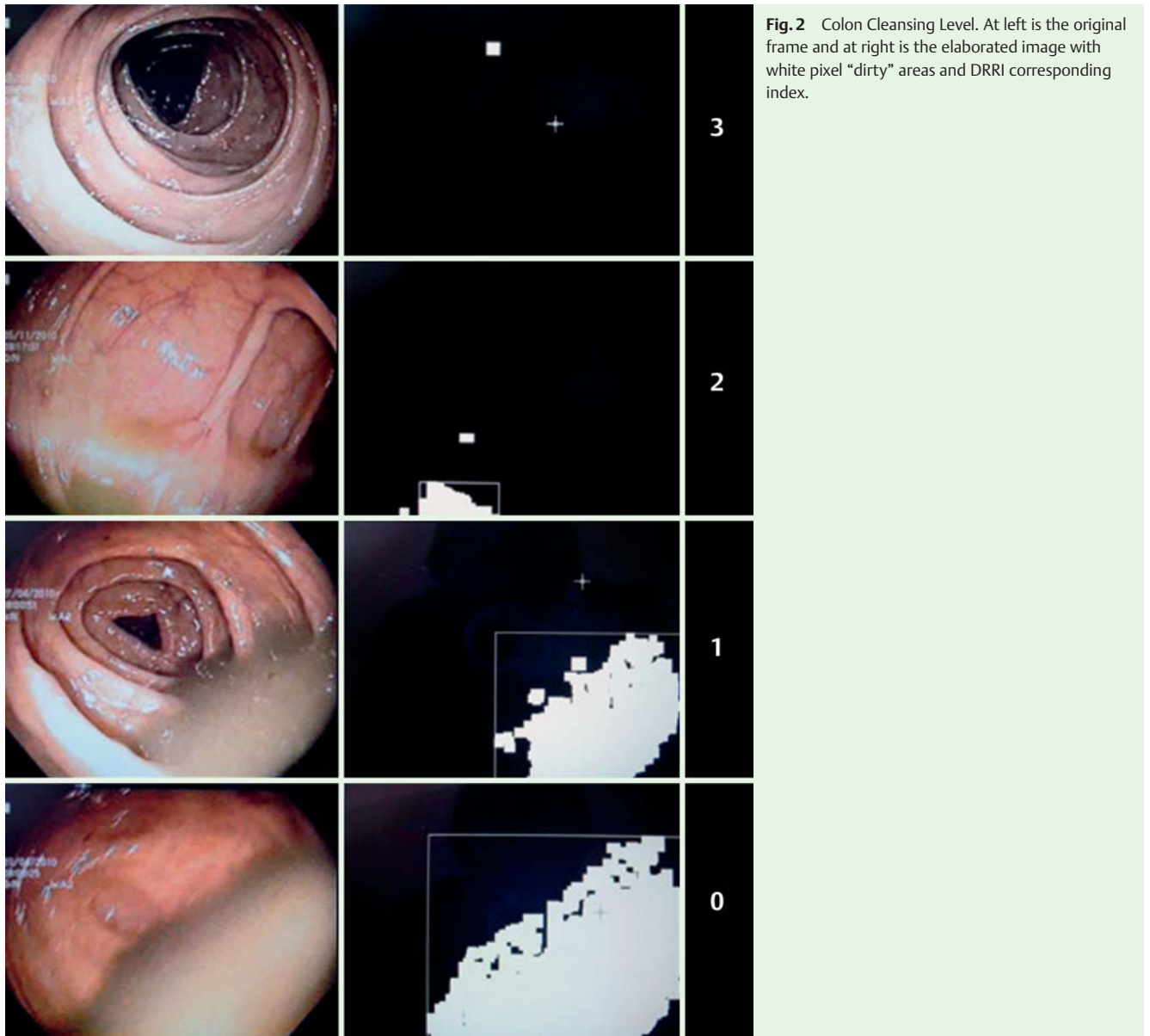

Fig. 3 Elaboration Data. The mean colon cleansing level is calculated by the elaboration of all frames.
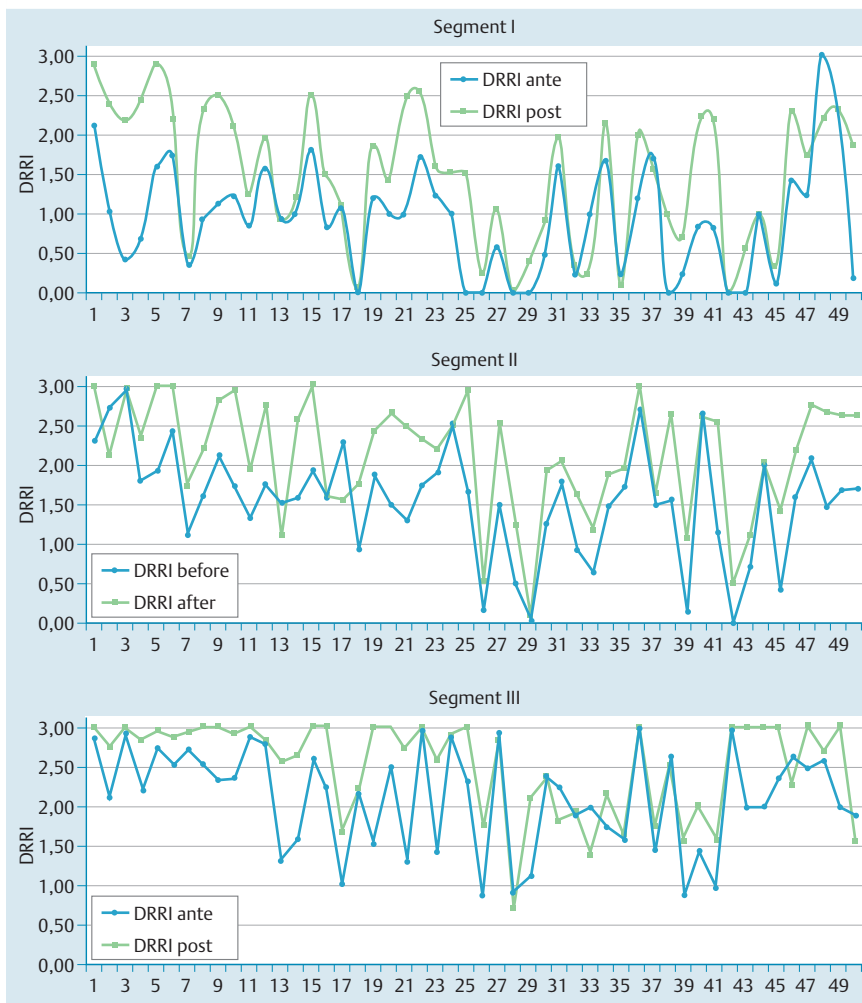
Fig. 4 Cleanliness of each colonoscopy divided by segment

pies and procedures with melanosis coli were excluded. The endoscopist started recording during endoscope withdrawal from the cecum. The colon was divided into three segments: cecum, ascending colon and hepatic flexure (first segment); transverse colon and descending colon (second segment); and the rectosigmoid segment (third segment). Hepatic flexure was identified by the liver impression on colon, whereas the rectosigmoid segment was 35 cm from the anal verge. Every segment was recorded twice, once before and once after colon cleansing with water injection and aspiration of the residual feces in such a way as to achieve the best possible cleansing of the mucosa. Clean colon segments were recorded twice without mucosal cleaning. Every segment was analyzed separately by our software. A T test for paired and unpaired data was used to define statistical significance.

Comparison between CCSP and Boston Bowel Preparation Scale

▼

A blind evaluation of the aggregated 300 truncated colonoscopy videos was performed by four endoscopists from three different endoscopy units. The validated Boston Bowel Preparation scale (BBPS) [7] was chosen to establish colon cleansing. Before evaluating the colonoscopy videos, each endoscopist viewed a 15-minute digital training video (available on http://www.cori.org/bbps/instruction.php) twice to enhance comprehension of BBPS.

Each endoscopist visualized all of the colonoscopy videos over a period of 3 days and gave each video a score from 0 to 3. We considered DRRI as the evaluation score of a fifth endoscopist and rounded up every score to the nearest integer. To assess interobserver reliability, the interclass correlation coefficient (ICC) was calculated of the scores after viewing of the colonoscopies and from the software's DRRI, using the methods of Shrout and Fleiss [11]. The strength of agreement was considered "Very Good" for ICC values between 0.81 and 1, "Good" between 0.61 and 0.80, "Moderate" between 0.41 and 0.60, "Fair" between 0.21 and 0.40, and "Poor" at <0.20 according to Altman DG [12]

Results

▼

According to the CCSP system (○ **Fig. 4**), an improvement in cleansing was seen in all colon segments after lavage. The overall average (SD) DRRI score for pre-lavage colonoscopies was 1.56 ± 0.52 versus 2.08 ± 0.59 after lavage ($P < 0.001$), for an approximate 33.3% improvement in cleansing after post-lavage (○ **Table 1**). The mean DRRI for the first segment was 0.99 ± 0.69 pre-lavage and of 1.80 ± 0.80 post-lavage, demonstrating an improvement in cleansing of 88.1% ($P < 0,001$). The DRRI for the second segment was 1.55 ± 0.70 pre-lavage and 2.13 ± 0.72 post-lavage, demonstrating an improvement in cleansing of 37.1% ($P < 0.001$). For the third segment, the DRRI was 2.07 ± 0.71 pre-lavage and 2.46 ± 0.56 post-lavage, for an improvement in cleansing of 18.8%

		Mean DRRI	SD	P value	% improvement
First segment	pre	0.99	0.69	<.001	81.8
	post	1.80	0.80		
Second segment	pre	1.55	0.70	<.001	37.1
	post	2.13	0.72		
Third segment	pre	2.07	0.71	<.001	18.8
	post	2.46	0.56		
Overall	pre	1.56	0.52	<.001	33.3
	post	2.08	0.59		

Table 1 Mean Pre- and Post-lavage colon cleansing

			Interval confidence 95%		
		ICC	Upper limit	Lower limit	P value
First segment	pre	0.89	0.83	0.93	.000
	post	0.90	0.85	0.93	.000
Second segment	pre	0.71	0.56	0.82	.000
	post	0.77	0.64	0.86	.000
Third segment	pre	0.77	0.64	0.86	.000
	post	0.80	0.69	0.88	.000
Overall	pre	0.77	0.64	0.86	.000
	post	0.82	0.72	0.88	.000

Table 2 Evaluation of interob-server agreement between endoscopists and CCSP using interclass correlation coefficient (ICC)

			Interval confidence 95%		
		ICC	Upper limit	Lower limit	P value
First endoscopist	pre	0.64	0.0	0.74	.000
	post	0.70	0.58	0.78	.000
Second endoscopist	pre	0.65	0.53	0.70	.000
	post	0.64	0.50	0.74	.000
Third endoscopist	pre	0.70	0.59	0.75	.000
	post	0.71	0.60	0.79	.000
Fourth endoscopist	pre	0.69	0.57	0.78	.000
	post	0.65	0.52	0.75	.000
Endoscopists' average score	pre	0.87	0.84	0.90	.000
	post	0.86	0.83	0.89	.000

Table 3 Evaluation of interob-server agreement between each endoscopist's BBPS score vs CCSP and the average BBPS score for all endoscopists vs CCSP using inter-class correlation coefficient (ICC).

($P<0,001$). Pre-lavage, the right colon segment was significantly dirtier (DRRI 0.99 [0.69]) than the left colon segment (DRRI third segment 2.07±0.71, $P<0.0001$ and DRRI second segment, 1.55± 0.70, $P<0.001$). A statistically significant difference in cleansing was recorded between the third and second segments (2.07± 0,71 vs 1.55±0.70, $P<0.001$).

As shown in ◗ **Table 2**, the overall interobserver agreement between endoscopists and the software pre-cleansing was 0.77 (95% CI, 0.64–0.86) and post-cleansing was 0.82 (95% CI, 0.72–0.88). The ICCs for the first segment pre- and post-cleansing were 0.89 (95% CI, 0,83–0.93) and 0.90 (95% CI, 0.85–0.94), for the second segment were 0.71 (95% CI, 0.56–0.82) and 0.77 (95% CI, 0.64–0.86) and for the third segment were 0.77 (95% CI, 0.64–0.86) and 0.80 (95% CI, 0.69–0.88). As explained in ◗ **Table 3**, we also calculated the interobserver agreement between each endoscopist and the software. For the first endoscopist vs CCSP, the ICC pre-cleansing was 0.64 (95% CI, 0.50–0.74) and post-cleansing was 0.70 (95% CI, 0.58–0.78); for the second endoscopist, it was 0.65 (95% CI, 0.53–0.70) pre-cleansing and 0.64 (95% CI, 0.50–0.74) post-cleansing; for the third endoscopist, the CCSP for pre-cleansing was 0.70 (95% CI, 0.59–0.75) and for post-cleansing 0.71 (95% CI, 0.60–0.79); and for the fourth endoscopist, the CCSP for pre-cleansing was 0.69 (95% CI,

0.57–0.78) and for post-cleansing it was 0.65 (95% CI, 0.52–0.75). We then calculated the average BBPS score across all four endoscopists and the interobserver correlation with DRRI score: pre-cleansing 0.87 (95% CI, 0.84–0.90) and post-cleansing 0.86 (95% CI, 0.83–0.89). In the histogram in ◗ **Fig. 5**, the ordinate indicates the number of videos and the abscissa indicates the three colon segments subdivided by the scores for each endoscopist/ CCSP and BBPS/DRRI evaluation.

Discussion
▼

As demonstrated by the very good interobserver reliability between the software and the average cleansing score for the endoscopists, CCSP has the potential to become a reliable and universal method to objectively assess colon cleanliness, replacing the endoscopist's subjective point of view.

A computer-assisted clean-colon scale could be an attractive alternative to the current proposed scales. With a system developed for clinical use, the software would be inserted into an endoscopy videoprocessor to objectively evaluate and quantify in colon cleanliness in real time. A DRRI established by the software in real time, while an exam is being carried out, could give
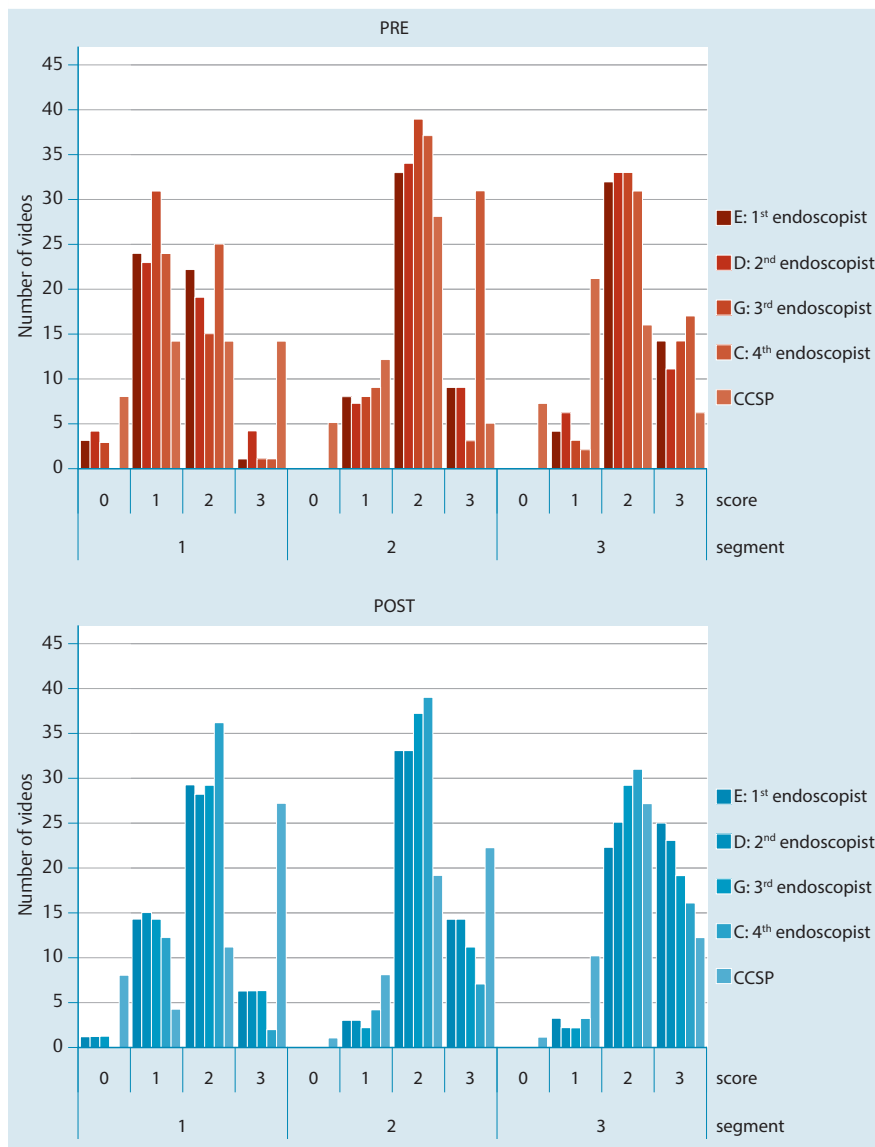
Fig. 5 Total number of videos for all endoscopists and CCSP subdivided by cleansing score and colon segment

the endoscopist an opportunity to compare his/her impression of colon cleanliness with that from the software.

"Flat" lesions and serrated polyps, usually reported in the right colon, can be missed if the mucosal surface is not clean enough [13 – 15]. It would be of particular diagnostic interest if the program makes it possible to identify a threshold of cleansing that would minimize missed lesions and improve colonoscopy quality as defined by completion rates, technique, and accuracy of inspection.

Furthermore, despite all the evidence, it remains difficult to select the best bowel preparation because many studies conducted in that areas are subject to bias. Thus, CCSP could be a good method to easily determine if one preparation is better than another, surpassing the old and intricate concept of interobserver and intraobserver agreement for establishing in clinical trials which of two bowel preparations results in the best cleansing.

During the development of the software, we struggled with the problem of melanosis coli, especially when the condition was severe. In fact, with it, the software recognizes the mucosa as dirty even when it is clean. We have not yet overcome that problem, but we are trying to recalibrate the software to account for this relatively frequent occurrence [16].

Another critical aspect of this method are the variables that can occur during the exam itself (e.g., polypectomy or biopsy). At those points, the video should focus on a particular zone of the colon that, from the viewpoint of cleanliness, is of interest only in regard to the index of the first frame sample in that zone. The software should automatically record the frames that follow and take them into consideration once the scope starts to move forward again.

In order to account for all possible variables found in clinical practice, both anatomic and procedural, an ample quantity of video data recorded by a large number of endoscopists using different equipment is needed so that the grading of bowel preparation can be defined and the software application modified as needed.

In conclusion, our preliminary data show that the Clean Colon Software Program is able to distinguish between clean and non-clean colon tracts and it is comparable to the validated Boston Bowel Preparation scale. To our knowledge, this is the first objective method of establishing colon cleansing that has been described in the literature.

Competing interests: None

Institutions

[1] Dpt of Hospital and Territory, Gastroenterology Unit, St Anthony Hospital, Padua
[2] Dpt of Medicine, Oncology Unit, San Donà di Piave, Venice
[3] Dpt of Medicine, Chioggia Hospital, Venice
[4] Dpt of Medicine, Gastroenterology Unit, San Bassiano Hospital, Bassano del Grappa, Vicenza
[5] Dpt of Medicine, Endoscopy Unit, Dolo Hospital
[6] Division of Endoscopy, European Institute of Oncology, Milan
[7] Dept. of Molecular Medicine, Laboratory of Public Health and Population Studies, University of Padua, Padua, Italy

References

1 *Puckett J, Soop M.* Optimizing colonoscopy preparation: the role of dosage, timing and diet. Curr Opin Clin Nutr Metab Care 2012; 15: 499–504

2 *Shaukat A, Mongin SJ, Geisser MS* et al. Long-Term Mortality after Screening for Colorectal Cancer. N Engl J Med 2013; 369: 1106–1114

3 *Singh H, Nugent Z, Demers AA* et al. The reduction in colorectal cancer mortality after colonoscopy varies by site of the cancer. Gastroenterology 2010; 139: 1128–1137

4 *Rex DK, Imperiale TF, Latinovich DR* et al. Impact of bowel preparation on efficiency and cost of colonoscopy. Am J Gastroenterol 2002; 97: 1696–1700

5 *Aronchick CA, Lipshutz WH, Wright SH* et al. Validation of an instrument to assess colon cleansing. [abstract]. Am J Gastroenterol 1999; 94: 2667

6 *Rostom A, Jolicoeur E.* Validation of a new scale for the assessment of bowel preparation quality. Gastrointest Endosc 2004; 59: 482–486

7 *Calderwood AH, Jacobson BC.* Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc 2010; 72: 686–692

8 *Halphen M, Hershback D, Gruss HJ* 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

9 *Gerard DP, Foster DB, Raiser MW* et al. Validation of a New Bowel Preparation Scale for Measuring Colon Cleansing for Colonoscopy: The Chicago Bowel Preparation Scale. Clin Transl Gastroenterol 2013; 4: e43 DOI 10.1038/ctg.2013.16

10 *Hassan C, Bretthauer M, Kaminski MF* et al. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy ESGE) Guideline. Endoscopy 2013; 45: 142–150

11 *Shrout PE, Fleiss JL.* Interclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420–428

12 *Altman DG.* Statistics for medical research. Chapman & Hall/CRC press; 1990

13 *Soetikno RM, Kaltenbach T, Rouse RV* et al. Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. JAMA 2008; 299: 1027–1035

14 *Kahi CJ, Hewett DG, Norton DL* et al. Prevalence and variable detection of proximal colon serrated polyps during screening colonoscopy. Clin Gastroenterol Hepatol 2011; 9: 42–46

15 *Brenner H, Chang-Claude J, Seiler CM* et al. Interval cancers after negative colonoscopy: population-based case-control study. Gut 2012; 61: 1576–1582

16 *Koskela E, Kulju T, Collan Y.* Melanosis coli. Prevalence, distribution, and histologic features in 200 consecutive autopsies at Kuopio University Central Hospital. Dis Colon Rectum 1989; 32: 235–239