Thieme

Endoscopy International Open

Original Article

# Automatic Evaluation of Small Bowel Cleanliness in Capsule Endoscopy: A Comparative Study of Four Preparations using Artificial Intelligence

Andrea Nevarez[1], Reinier Noorda[2], Noelia Alonso-Lazaro[1], Cristina Carretero[3], César Prieto de Frías[3], Begoña González Suárez[4], Antonio Giordano[5], Victoria-Alejandra Jiménez-García[6], Laura Lorenzo Gonzalez[6], Ángel Caunedo-Álvarez[6], Julio Valle Muñoz[7], Almudena García Vela[7], Marisol Lujan-Sanchís[8], Lucía Ruiz Sánchez[8], Francisco Sanchez Ceballos[9], Juan Luis Mendoza Hernández[9], Javier García Lledó[10], Oscar Nogales Rincón[11], Juan Egea Valenzuela[12], María José Mesa López[12], Ainara Merino Zubizarreta[13], Pamela Estévez Boullosa[14]†, Alfonso Martinez-Turnes[15], Marco Bustamante-Balén[1], Valery Naranjo[16], Vicente Pons Beltrán[1]

Affiliation addresses are listed at the end of the article.

## ABSTRACT

**Background** Diagnostic yield (DY) of small bowel capsule endoscopy (SBCE) is limited by the presence of intestinal content. An objective assessment of the cleanliness grade of the small bowel (SB) is currently lacking. We developed and validated an artificial intelligence (AI) tool to assess SB cleanliness in a multicenter observational study comparing four types of bowel preparations.

**Objectives** Identify the best preparation for SBCE applying a system based on convolutional neural networks that assess the SB degree of cleanliness.

**Design** We included 372 SBCE consecutive procedures with four different bowel preparations. Group A (*n* = 93): liquid diet (LD), Group B (*n* = 92): macrogol 3350 + ascorbic acid 1000 mL, Group C (*n* = 94): sodium picosulfate, and Group D (*n* = 93): macrogol 500 mL. The degree of cleanliness of the SB was evaluated with the automatic tool (0–100%). Gastric and SB transit times, DY, and degree of satisfaction with the preparation were also evaluated.

**Results** Mean age of included patients was 58.14 years (confidence interval 95% 18–93 years); 55.9% were women. There were no significant differences in the degree of cleanliness of the SB evaluated with AI between groups (Mean Group A: 51.28%; B: 54.79%; C: 52.44%; D: 55.34%; p=0.415), nor in transit times or DY. LD was the best-tolerated preparation. Multivariable analysis identified age and diabetes as independent predictors of lower cleanliness.

**Conclusion** We demonstrated using an AI-based tool, that LD provides adequate preparation for SBCE, questioning the routine administration of laxatives and supporting a patient-centered approach.

**Keywords** endoscopy small bowel, capsule endoscopy, small bowel endoscopy, quality and logistical aspects, image and data processing, documentation

**received** June 05, 2025 | **accepted after revision** June 16, 2026 | **article published online** 2026

**Bibliography** Endosc Int Open 2026; a28974847 **DOI** 10.1055/a-2897-4847 **Art ID** EIO-2025-06-3829-OA


† Deceased.

**Correspondence** Vicente Pons Beltrán, Department of Gastroenterology, La Fe Health Research Institute – IIS La Fe, Valencia, Valencian Community, Spain, Email: pons_vicbel@gva.es

## Background

Capsule endoscopy (CE) is an essential tool for diagnosing small bowel (SB) diseases. CE manufacturers recommend a liquid diet (LD) and 12 hours of fasting on the day before CE ingestion. However, the diagnostic yield (DY) is limited due to the presence of intestinal content, which prevents adequate visualization of the mucosa, resulting in poor-quality images.[1] American and European guidelines suggest some benefits of laxative use in SB CE (SBCE)[2,3]; nonetheless, conclusions related to the time of administration of the laxative or finalization rates are lacking.

Research papers comparing different types of bowel preparations for SBCE have used numerous scales. Several of them have

not been validated, are complex to use, and furthermore, the interobserver agreement is low, as demonstrated by Dray et al.[4] in a recent publication.

In the past few years, computer-dependent scales have been implemented for the evaluation of SB cleanliness, and new technologies based on artificial intelligence (AI) have also been used for this purpose. However, these scales have not been widely used or applied in large-scale studies. Most AI tools that have been explored are for automatic detection, identification, or characterization of SB lesions, becoming increasingly important to publish in a real-world setting a model of AI in CE.[5]

Our group developed an AI tool that evaluates the degree of cleanliness in SBCE PillCam systems, which automatically detects the proportion of regions with intestinal content in CE images, providing a cleanliness score presented as a percentage in each image and video analyzed. This system achieved an accuracy of 95.23% (sensitivity of 96.18% and specificity of 94.33%) with an acceptable interobserver agreement (intraclass coefficient 0.86–0.90).[6]

Our work aimed to use the designed AI tool in a multicenter observational study comparing three different types of laxatives with LD administered before SBCE procedures.

# Design

## Study design

This was an observational, prospective, multicenter trial. Patients were assigned based in the standard clinical practice of the participating hospitals into the following groups: A: 4 l no volume-restricted clear LD 24 hours before CE ingestion; Group B: Moviprep Norgine BV Spain (macrogol 3350 + ascorbic acid 1000 mL); Group C: Citrafleet Casen Recordati (sodium picosulfate); and Group D: Pleinvue Norgine BV Spain (macrogol 3350 500 mL). Laxatives were administered 2 hours before SBCE. CE was ingested with water and 200 mg (2 mL) of simethicone. An LD was initiated 2 h after CE administration, followed by a solid diet 4 hours after. No prokinetics were used in this trial. All hospitals participating in the study used the PillCam SB3 system. Data were collected using the REDCap database.[7,8]

## Study Population

From April 2020 to January 2022, 372 patients who underwent SBCE at 11 hospitals in Spain were included. Exclusion criteria were patients with hepatic, cardiovascular, or renal diseases that prevented the use of laxatives; intestinal obstruction; hypersensitivity to any of the laxative components; and age < 18 years. Proper indication for CE was not evaluated in this study.

## CE Video Analysis

All videos included in the study were anonymized and subsequently evaluated for findings by an independent expert at each participating hospital. CE findings were reported using the CE Structured Terminology and classified as relevant or nonrelevant. Each video was later analyzed by an automatic system based on a novel convolutional neural network (CNN), consisting of four blocks of convolutional layers followed by batch normalization layers, which were then followed by dense layers. In all convolutional and dense layers, LeakyReLU was used as the activation function.

All included videos were analyzed by the AI tool. This tool, named hereafter Endoclean, evaluated the proportion of intestinal content obtained in all included images and later provided a proportion score that represents the percentage of cleanliness in each video included in our study. An example of frames with different intestinal contents evaluated using Endoclean is shown in **Fig. 1**.

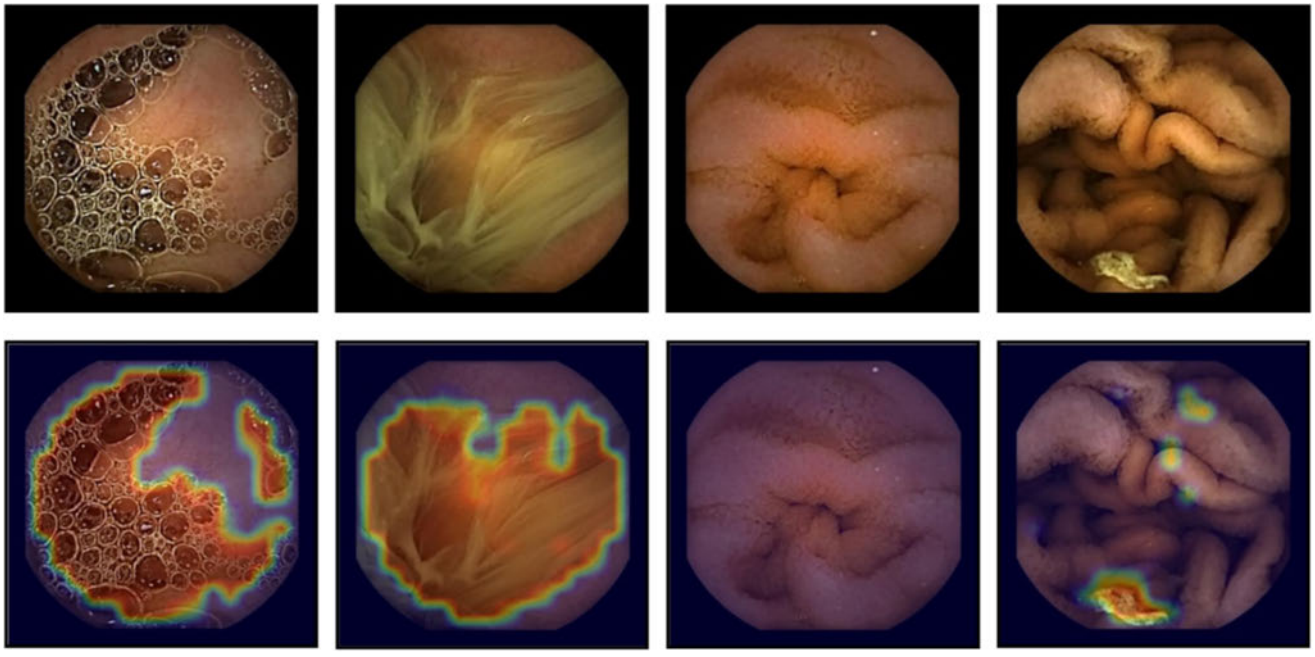

**Fig. 1** Example frames of CE with different content evaluated by Endoclean.

### Primary and Secondary Endpoints

The primary endpoint was to evaluate the degree of cleanliness in SBCE videos using the AI-based tool comparing four types of preparations.

Gastric transit time (GTT), SB transit time (SBTT), completion rate, and tolerance among the different preparations were considered secondary endpoints.

DY, incomplete procedures, and capsule retention were also evaluated.

### Definitions

GTT: time between the first gastric image and the first duodenal image.

SBTT: time between the first duodenal image and the first cecal image or the last SB image.

Completion rate: capsule passing through the ileocecal valve or entering the colon.

DY: relevant diagnoses that were considered likely to explain the cause of symptoms in relation to the main indications.

Incomplete procedures: capsules that did not reach the cecum.

Capsule retention: capsules that remain in the digestive tract for a minimum of 2 weeks.

### Statistics

The sample size was based on a clinically relevant difference of 20% in the proportion of adequate SB cleansing between the groups (from 50% to 70%).[9] Assuming a significance level of 5% ($\alpha = 0.05$) with a statistical power of 80% ($\beta = 0.20$), it was calculated that 93 patients per group would be required to detect this difference using a two-sided test for the comparison of independent proportions.

Continuous variables are presented as means and standard deviations and were analyzed using the Kruskal–Wallis and ANOVA tests. Categorical variables are presented as frequencies and percentages and were analyzed using the chi-square test.

Multivariable linear regression model was conducted for the primary outcome (degree of SB cleanliness). Additional regression analyses were performed for secondary outcomes where appropriate. Regression coefficients with standard errors and *p*-values were reported.

For clinical explanation, the continuous Endoclean score was categorized into four qualitative levels: “excellent” >75%, “very good” 50–75%, “good” 49–25%, and “poor” <25%. A further analysis clustering groups with <25% cleanliness (“poor”) and those with >25% cleanliness (“good”) was also executed and compared laxatives as a group versus LD.

Statistical analyses were performed using SPSS v29.0.0.0. A two-sided *p*-value of <0.05 was considered statistically significant.

### Ethical Aspects

All patients provided informed consent, and the study protocol followed the ethical guidelines of the Declaration of Helsinki. This protocol was approved by the Ethical Committee of each of the institutions that participated in the study and was additionally assessed by the Spanish Agency of Medications and Sanitary Products, approved later on April 10, 2019 (code VPB-2019-01).

## Results

A total of 372 patients were included: 93 in Group A, 92 in Group B, 94 in Group C, and 93 in Group D. Patients had a mean age of 58.14 years (SD ±18.29) and 208 (55.9%) were female. Personal history of hypertension (36.3%) and diabetes (20.7%) were the most common comorbidities (**Table 1**).

The two main indications for SBCE were SB bleeding (54.9%) and chronic diarrhea (20.7%) (**Table 1**).

For key covariates such as age, diabetes, preparation type, and overt bleeding, multivariable linear regression analysis demonstrated that the type of bowel preparation was not independently associated with higher Endoclean cleanliness scores. Older age ($\beta = -0.175$, $p = 0.004$) and diabetes ($\beta = -6.359$, $p = 0.014$) were associated with lower cleanliness scores (**Table 2**).

### Findings and Diagnosis of SBCE

The most common findings were vascular lesions and erosions, with no statistically significant differences between groups (**Table 3**).

### Degree of Cleanliness of the SB

The mean degree of cleanliness represented as a percentage score was 51.28% in Group A, 54.79% in Group B, 52.44% in Group C, and 55.34% in Group D. There were no statistically significant differences in the mean degree of cleanliness between the groups, as seen in **Table 4** and **Fig. 2**. Degree of cleanliness was not assessed per SB segments.

When bowel cleanliness was analyzed using the qualitative Endoclean categories, a significant difference between preparation groups was observed ($\chi^2$ test, $p = 0.003$). This difference was mainly determined by the proportion of examinations classified as “good.” The distribution of higher-quality categories (“excellent” and “very good”) was similar across groups (**Table 5**).

The mean proportion of videos with cleanliness over 75% and less than 24% were also analyzed, and no differences were noticed ($p = 0.231$ vs. $p = 0.064$, respectively). However, more videos with poor quality (less than 24% cleanliness) were observed in the laxative group versus LD group (27 vs. 2 $p = 0.064$).

When categories were dichotomized into “good” versus “poor,” no significant differences were observed between groups ($p = 0.123$). Further analysis was also performed by grouping cleanliness categories into “good” and “poor” and laxatives compared as a group versus LD alone, showing a tendency toward higher proportion of “poor” examinations compared to LD (9.7% vs. 2.1%, $p = 0.019$).

We performed a multivariate linear regression model to evaluate the impact of transit times, preparation type and proportion of bowel cleanliness, using the Endoclean cleanliness score as a continuous-dependent variable. SBTT was identified as a significant independent predictor ($R = 0.028$; $p = 0.019$), where slower transit was associated with lower cleanliness levels. The type of preparation did not significantly influence cleanliness scores ($p =$ >0.157 for all groups).

**Table 1** Characteristics and indications for SBCE in each group.

| Characteristics | Group A Liquid Diet *N* = 93 | Group B Moviprep *N* = 92 | Group C Citrafleet *N* = 94 | Group D Pleinvue *N* = 93 | Total *N* = 372 | *p*-Value |
|---|---|---|---|---|---|---|
| Female/Male (%) | 50/43 (53.76/46.24) | 51/41 (55.43/44.57) | 57/37 (60.64/39.36) | 50/43 (53.76/44.24) | 208/164 (55.91/44.09) | NS |
| Mean age Female/Male (±SD) | 60.32 (±17.93)/ 62.51 (±18.44) | 56.16 (±17.98)/ 58.05 (±18.02) | 61.11 (±18.71)/ 60.35 (±18.57) | 51.68 (±16.69)/ 55.35 (±18.79) | 57.44 (±18.14)/ 59.03 (±18.43) | **^0.012** |
| Hospitalized/nonhospitalized | 8/85 (91.40%) | 4/88 (95.65%) | 5/89 (94.68%) | 5/88 (94.62%) | 22/350 (94.09%) | NS |
| Hypertension | 37 (39.78%) | 29 (31.52%) | 41 (43.62%) | 28 (30.11%) | 135 (36.29%) | NS |
| Diabetes | 26 (27.96%) | 10 (10.87%) | 23 (24.47%%) | 18 (19.35%%) | 77 (20.70%) | **0.025** |
| **SBCE indications (%)** | | | | | | |
| Small bowel bleeding | 62 (66.66%) | 54 (58.69%) | 66 (70.21%) | 41 (44.08%) | 223 (59.94%) | NS |
| Chronic diarrhea | 17 (18.28%) | 21 (22.83%) | 21 (22.34%) | 18 (19.35%) | 77 (20.70%) | NS |
| Suspected inflammatory bowel disease | 21 (22.58%) | 17 (18.48%) | 15 (15.96%) | 22 (23.66%) | 75 (20.16%) | NS |
| Abdominal pain | 18 (19.35%) | 11 (11.96%) | 22 (23.40%) | 13 (13.98%) | 64 (17.20%) | NS |

NS: Not significant.
SD: Standard Deviation.
^ Difference in age with younger patients in Group D.

**Table 2** Multivariable analysis of factors influencing cleanliness.

| Predictor | *B* | *SE* | 95% CI (LL;UL) | *t* | *p* |
|---|---|---|---|---|---|
| (Intercept)[a] | 65.331 | 4.121 | 57.22; 73.43 | 15.851 | ***<0.001*** |
| Age | −0.175 | 0.060 | −0.293; −0.056 | −2.908 | ***0.004*** |
| Diabetes (yes) | −6.359 | 2.584 | −11.44; −1.27 | −2.461 | ***0.014*** |
| **Preparation type** | | | | | |
| Group B vs. Group A | 1.982 | 2.740 | −3.40; 7.37 | 0.723 | 0.470 |
| Group C vs. Group A | 0.775 | 2.710 | −4.55; 6.10 | 0.286 | 0.775 |
| Group D vs. Group A | 1.868 | 2.730 | −3.50; 7.23 | 0.684 | 0.494 |
| **Clinical Indications Overt bleeding** | | | | | |
| Melena | −6.121 | 3.687 | −13.37; 1.13 | −1.660 | 0.098 |
| Hematochezia | −2.589 | 7.164 | −16.67; 11.50 | -0.361 | 0.718 |
| Rectal bleeding | −6.638 | 4.309 | −15.11; 1.83 | −1.540 | 0.124 |
| FOBT (positive) | 0.746 | 4.984 | −9.05; 10.54 | 0.150 | 0.881 |

[a]Reference level; *B* = unstandardized coefficient; *SE* = standard error; CI = confidence interval; *p* = *p*-value; OBT = fecal occult blood test. Group A (Liquid Diet) is the reference category for preparation type. Significant predictors ($p < 0.05$) are shown in bold.

## Diagnostic Yield

Findings were observed in 73.3% of explorations and 26.07% were considered as relevant. No differences between groups were found regarding to the DY when we compared relevant diagnosis with the cause of symptoms and main indications.

## Transit Time in SBCE

No differences were observed between groups regarding gastric and SB times: GTT mean time was 41.82 min in Group A, 41.65 min in Group B, 52.76 min in Group C, and 49.79 min in Group D ($p = 0.322$). SBTT mean time was 254.52 min in Group A, 272.32 min in Group B, 262.77 min in Group C, and 239.36 in Group D ($p = 0.212$).

## Incomplete procedures, retentions, and hospitalized patients

In our cohort of patients, 19 (5%) SBCE were incomplete; 10 of them belonged to Group D (Pleinvue) ($p = 0.01$). Capsule

**Table 3** Findings in SBCE groups.

| Findings in SBCE | Group A Liquid Diet *N* = 93 | Group B Moviprep *N* = 92 | Group C Citrafleet *N* = 94 | Group D Pleinvue *N* = 93 | Total *N* = 372 | *p*-Value |
|---|---|---|---|---|---|---|
| Vascular lesions | 25 (26.9%) | 27 (29.3%) | 16 (17%) | 15 (16.1%) | 83 (22.3%) | NS |
| Erosions | 20 (21.5%) | 16 (17.4%) | 18 (19.15) | 25 (26.9%) | 79 (21.2%) | NS |
| Mucosal changes | 18 (19.4%) | 14 (15.2%) | 24 (25.5%) | 20 (21.5%) | 76 (20.4%) | NS |
| Ulcers | 9 (9.7%) | 8 (8.7%) | 13 (13.8%) | 19 (20.4%) | 49 (13.2%) | NS |
| Normal SBCE | 23 (24.7%) | 22 (23.9%) | 28 (29.8%) | 24 (25.8%) | 97 (26.1%) | NS |

NS: Not significant.

**Table 4** Mean degree of cleanliness in different groups analyzed using Endoclean tool.

| Degree of cleanliness | Group A Liquid Diet *N* = 93 | Group B Moviprep *N* = 92 | Group C Citrafleet *N* = 94 | Group D Pleinvue *N* = 93 | *p*-Value |
|---|---|---|---|---|---|
| Lowest value | 10 | 3 | 7 | 9 | |
| Highest value | 82 | 98 | 88 | 96 | |
| Mean | 51.28 (CI 48.33–54.23) | 54.79 (CI 50.23–59.35) | 52.44 (CI 48.82–56.05) | 55.34 (CI 50.99–59.69) | 0.415 |
| Standard deviation | 14.304 | 22.026 | 17.656 | 21.121 | |

CI Confidence intervals

retention was observed in eight patients: four of them in Group B (Moviprep) and four in Group D (Pleinvue). Among incomplete capsule examinations and capsule retention, no clear differences in bowel cleanliness categories were observed between preparation groups ($p = 0.112$ vs. $p = 0.983$, respectively).

In incomplete procedures, no significant differences were observed regarding prior abdominal surgery ($p = 0.219$), age ($p = 0.355$), diabetes mellitus ($p = 0.229$), or bedridden status ($p = 0.382$).

Twenty-two patients were hospitalized (5.9%) at the time of the procedure, and only two of these patients had incomplete examinations (one in the LD group and the other in the Pleinvue group). The use of laxatives did not obtain better SB preparation in this group of patients ($p = 0.112$).

### Tolerability of the Procedure

Tolerance to the various preparations used was evaluated through a survey that had been previously validated in our earlier research.[9] Best-tolerated preparation was found in Groups A (LD) and C (Citrafleet) ($p < 0.001$), and the least-tolerated preparations were observed in Groups B (Moviprep) and D (Pleinvue) ($p = 0.05$). **Table 6** shows the global experience with the preparations, with more patients in Groups A and C reporting excellent and good experiences, respectively. No significant differences in the degree of cleanliness of the SB were observed between those patients with excellent experience and those with better or worse experience. All patients completed bowel preparation.

A significant difference in patient-reported tolerability was observed between preparation groups ($\chi^2$ test, $p = 0.001$). The LD group showed a higher proportion of excellent tolerance, whereas the Pleinvue group showed a higher proportion of normal tolerance.

## Discussion

In this study, we have shown that the degree of cleanliness did not significantly change depending on the type or preparation (Group A: 51.58%; Group B: 54.79%; Group C: 52.44%, and Group D: 55.34). No significant differences were observed between groups in terms of DY, GTT, and SBTT.

Despite technological advances in SBCE, preparation and assessment of adequate quality of cleanliness in SB remains an important unresolved issue. Quality of SB visualization (QV) and DY can be affected by the presence of fluids, residues, bubbles, and intestinal content, mostly when diagnosing vascular lesions or angiodysplasia.[1]

European and North American guidelines recognize that the use of bowel preparation improves mucosal visualization.[3,10] Multiple clinical reports have also described the use of laxatives in SBCE[11–13]; however, there is still no recommendation for the use of a specific preparation, dose, or time of administration.[13] The absence of an objective and validated scale with the least possible variability plays an important role in this matter.[14] Frequently used operator-dependent scales, such as the scale proposed by Brotz, have fair to moderate interobserver reproducibility.[4] In addition, Cortegoso et al.[15] also performed a systematic review and meta-analysis to evaluate agreement in Video CE reading, observing heterogeneity in the revised literature and a suboptimal inter/intraobserver agreement. Also suggesting that the use of AI tools could offer an opportunity to standardized observer-independent evaluation.

The development of automatic scales has been previously described. The initial attempt conducted by Van Weyenberg et al.,[16] measured the differences between colors in the tissue color bar as observed in the software reading SBCE system, yielding promising results. AI systems have also been employed to assess cleanliness in the SB. Leenhardt et al.[17] developed an

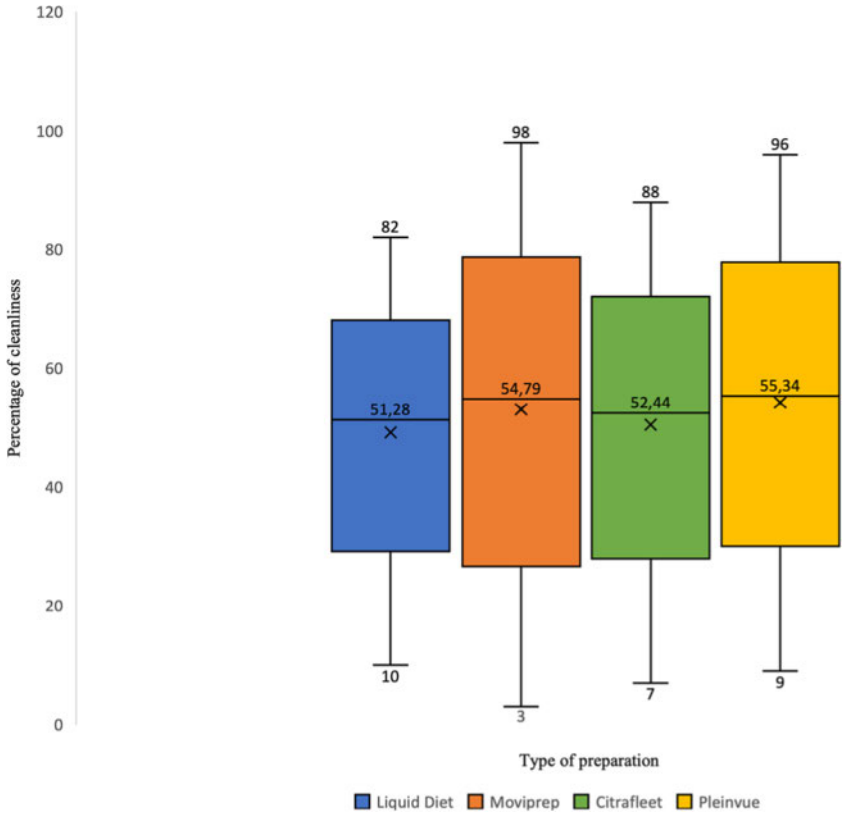


**Fig. 2** Small bowel cleanliness scores by preparation group. The boxplot displays the AI-calculated cleanliness percentage (0–100%) for each group analyzed by Endoclean. Boxes represent the interquartile range (IQR), with the median indicated by the horizontal line and the mean by the "x" symbol. To ensure clarity across all renders, the color coding is as follows: Dark Blue = Liquid Diet (Group A); Orange = Moviprep (Group B); Dark Green = Citrafleet (Group C); and Light Blue = Pleinvue (Group D). No statistically significant differences were observed between groups ($p = 0.415$).

**Table 5** Contingency table with qualitative categories.

| Endoclean. Qualitative categories | Group A Liquid Diet *N* = 93 | Group B Moviprep *N* = 92 | Group C Citrafleet *N* = 94 | Group D Pleinvue *N* = 93 | Total |
|---|---|---|---|---|---|
| Excellent | 7 | 19 | 8 | 17 | 51 |
| Very good | 39 | 38 | 50 | 41 | 168 |
| Good | 45* | 26 | 26 | 27 | 124 |
| Poor | 2 | 9 | 10 | 8 | 29 |

***X*² *test***

| | $X^2$ | df | *p* |
|---|---|---|---|
| $X^2$ | 24.8 | 9 | 0.003 |
| *N* | *372* | | |

*Post hoc analysis showed that the difference was mainly driven by a higher proportion of "good" cleanliness in the liquid diet group.

**Table 6** Global experience between preparations.

| Global experience with preparation | Group A Liquid Diet *N* = 93 | Group B Moviprep *N* = 92 | Group C Citrafleet *N* = 94 | Group D Pleinvue® *N* = 93 | Total |
|---|---|---|---|---|---|
| Excellent | **18** | 7 | 9 | 6 | 40 |
| Good | **49** | 43 | **53** | 30 | 175 |
| Normal | 26 | 42 | 31 | **56** | 155 |

***X*² *test***

| | $X^2$ | df | *p* |
|---|---|---|---|
| $X^2$ | *29.6* | *6* | <0.001 |
| *N* | *370* | | |

algorithm based on neural networks with highly sensitive results (sensitivity 90.3%, specificity 83.3%, and accuracy 89.7%). Nam et al.[18] achieved an accuracy of 93% (sensitivity and specificity of 81%). While Houdeville et al.[19] designed an algorithm that measures the number of bubbles in SBCE recordings.

Using our AI automatic system, no statistically significant difference between preparations was observed. No differences were found in the proportion of procedures with better quality (more than 75% cleanliness) and those with low quality (less than 24% cleanliness) between groups. We also observed more procedures in the laxative group with a lower degree of cleanliness than in the LD group (27 vs. 2).

These results suggest that laxative use does not enhance visualization quality (QV) of SB. This is contrary to the work recently presented by Rahmi et al.[20] where the use of Polietilenglicol improved the cleanliness of the SB. However, Rahmi's work applied a subjective scale to analyze QV, unlike our work that used AI.

Even though the mean cleanliness scores observed in our study (approximately 52–55%) may appear low, it is important to mention that Endoclean generates a continuous quantitative score representing the proportion of frames with adequate mucosal visualization throughout the CE video. Unlike traditional operator-dependent scales, this value does not correspond to predefined categorical grades of bowel cleanliness.

Currently, there is no validated cutoff to define adequacy using AI-based cleanliness scores in CE. Therefore, establishing clinically meaningful thresholds remains an important objective for future research. Consequently, the cleanliness values reported in this study should be interpreted as a quantitative measure of mucosal visualization rather than a categorical assessment of bowel preparation quality. A score of approximately 50% does not necessarily correspond to inadequate visualization in clinical practice but rather reflects the strict frame-by-frame classification performed by the CNN.

In addition to the continuous quantitative analysis, a qualitative categorization of bowel cleanliness ("excellent," "very good," "good," "poor") based on the Endoclean score was performed in our study to facilitate clinical interpretation and demonstrated a global difference between preparation groups ($\chi^2$ test, $p = 0.003$), mainly driven by variations in the proportion of examinations classified as "good" in the LD group. While the distribution of higher-quality categories ("excellent" and "very good"), which are more clinically relevant, was comparable across groups.

When laxative-based preparations were analyzed as a single group, a higher proportion of poor-quality examinations was observed compared with LD.

Moreover, we detected with a multivariable linear regression that SBTT is an independent predictor of cleanliness ($p = 0.019$), with slower transit associated with poorer visualization. These findings confirm that as transit slows down, cleanliness scores significantly decrease and suggest that clinical focus should remain on managing patient factors like diabetes and age, rather than the timing of the laxatives seen in our cohort. However, future research could be conducted in this regard.

In multivariable linear regression analysis adjusting for relevant clinical variables (age, presence of diabetes, clinical indication), no independent association with SB cleanliness was seen. Compared with LD, none of the laxative-based preparations showed a statistically significant difference in cleanliness (Preparation B vs. A: β 1.982, $p = 0.470$; Preparation C vs. A: β 0.775, $p = 0.775$; Preparation D vs. A: β 1.868, $p = 0.494$).

Although factors such as older age and diabetes were associated with slightly lower cleanliness scores (consistent with previous literature), the type of preparation did not independently influence the quality of visualization.

Even though these adjustments reduce confounding, residual bias inherent to the nonrandomized design cannot be completely excluded. These findings should be viewed as real-world evidence that warrants further confirmation through large-scale randomized controlled trials using AI-based tools.

Regarding DY, there were no differences between groups either. The use of laxatives, compared to LD, did not contribute to uncovering a more relevant diagnosis, regardless of the type of laxative used.

Likewise, there were no significant differences in GTT and SBTT between groups. This aligns with the results of the meta-analysis by Gkolfakis et al.[12] The use of laxatives does not reduce the time of SBCE or affect transit times. This was also evident in hospitalized patients included in our study (22/5.9%), where the use of laxatives did not influence transit times or the degree of cleanliness.

Additionally, incomplete studies (19/5%) and retentions (8/2.1%) identified within our patient cohort did not exhibit any significant differences in the degree of cleanliness between groups. Although a higher proportion of incomplete examinations was observed in the Pleinvue group, additional analysis did not demonstrate a clear association between bowel cleanliness and incompleteness. Given the relatively small number of these types of procedures, these findings should be taken with caution.

Differences in tolerability were observed between preparation strategies, with LD being better tolerated than laxative-based regimens, with Citrafleet being the most tolerated laxative. However, these findings should be interpreted cautiously, as tolerability was analyzed as a secondary exploratory outcome and no further analysis was made.

We believe that in patients in whom SBCE must be repeated due to lack of adequate visualization of the SB and who were previously prepared with an LD, an alternative bowel preparation with laxatives could be Citrafleet due to its better acceptance.

The main limitation of our work is its observational design. However, this approach was chosen to avoid changing the participating institutions' usual clinical practice, offering actual conditions of standard practice with the use of SBCE in a real-world scenario. The decision not to randomize was also influenced by the challenges associated with introducing laxatives not approved by the hospital of origin, due to regulatory and administrative constraints. Laxatives that were not approved locally could not be implanted. Consequently, randomization was not feasible. We also believe that the use of our AI algorithm in future randomized studies can help obtain more definitive results, eliminating possible biases.

Another important factor that may influence bowel cleanliness is the timing of the laxative administration. In our study, laxatives were administered 2 hours before capsule ingestion, according to the routine clinical practice of the participating centers. This approach differs from most published preparation regimens, in which purgatives are typically administered the evening before the examination or using split-dose protocols. Several studies and meta-analyses have suggested that the earlier administration of purgatives may improve mucosal visualization during CE. Marmo et al.[13] showed better cleansing when purgatives were given 1–6 hours before SBCE compared to longer intervals (laxatives administered after the ingestion of CE had the highest rate of adequate SB cleansing with an RD 0.33, administration between 1 and 6 h before SBCE had an RD 0.28, 6 to 12 h had an RD 0.21 and ≥12 h had an RD 0.05).

Furthermore, the 2025 umbrella meta-analysis by Lei et al.[21] concluded that same-day dosing provides the strongest benefit compared to the other regimes in cleansing quality (RR = 1.75; 95% confidence interval [CI], 1.33–2.29), followed by postingestion regimens (RR = 1.59; 95% CI, 1.17–2.16). Our findings indicate that even when using same-day timing, purgatives did not offer a significant advantage over LD.

Despite the theoretical benefit of bowel preparation administered shortly before capsule ingestion, no significant association was observed in our study. In contrast, SB slower transit time showed a significant correlation with cleanliness, suggesting that other patient-specific factors may partially influence mucosal visualization, especially those associated with age or the presence of diabetes.

Another limitation in our study was the global evaluation of the percentage of cleanliness at a video level without separate proximal and distal SB analysis. Segmental evaluation could provide further insights, particularly given the known differences in residue distribution along the SB.

We also included only those institutions where PillCam SB system was used, not assessing other types of CE, since the AI algorithm was only trained with this system. However, the use of AI to evaluate cleanliness using an objective scale adds consistency to our results, eliminating subjectivity and interobserver variability. Our tool can also help establish quality standards in SB preparation.

A potential concern with deep learning models applied to CE images is their sensitivity to textural features, which may theoretically introduce bias when different types of luminal content are present. As a CNN-based tool, we acknowledge that Endoclean may be sensitive to domain shifts; therefore, our conclusions apply to the acquisition conditions used in this study (PillCam SB3 with routine simethicone), and its performance may differ with other systems or markedly different contaminant appearances.

## Conclusions

In conclusion, this multicenter observational study, demonstrated, using an AI-based tool, that an LD provides cleanliness and DY comparable to those of laxative-based preparations for SBCE. In the adjusted analyses, laxative-based preparations were not independently associated with improved SB cleanliness compared with LD. They revealed that cleanliness is mainly driven by patient intrinsic factors (age, diabetes) rather than the choice of regime used.

Our findings question the necessity of routine laxative administration and support for a patient-centered efficient approach.

## Author Affiliations

1 Department of Gastroenterology, La Fe Health Research Institute – IIS La Fe, Valencia, Valencian Community, Spain
2 Human-Centered Technology Institute (HUMAN-tech), Universidad Politècnica de Valéncia, Valencia, Valencian Community, Spain
3 Department of Gastroenterology, Clinica Universidad de Navarra, Pamplona, Navarra, Spain
4 Servicio de Gastroenterología, Hospital Clínic de Barcelona ICMiD, Barcelona, Catalunya, Spain
5 Gastroenterologia, Hospital Clínic de Barcelona, Barcelona, Catalunya, Spain
6 Department of Gastroenterology, Hospital Universitario Virgen Macarena, Sevilla, Sevilla, Spain
7 Department of Gastroenterology, Complejo Hospitalario Universitario de Toledo, Toledo, CM, Spain
8 Department of Gastroenterology, Consorci Hospital General Universitari de Valencia, Valencia, Valencian Community, Spain
9 Department of Gastroenterology, Hospital Clínico San Carlos, Madrid, Comunidad de Madrid, Spain
10 Department of Gastroenterology, Hospital General Universitario Gregorio Marañón, Madrid, Community of Madrid, Spain
11 Department of Gastroenterology. Endoscopy Unit, Hospital General Universitario Gregorio Marañón, Madrid, Spain
12 Unidad de Gestión Clinica de Digestivo, Hospital Clínico Universitario Virgen de la Arrixaca, El Palmar, Región de Murcia, Spain
13 Department of Gastroenterology, Hospital de Basurto, Basurto-Zorroza, PV, Spain
14 Hospital Álvaro Cunqueiro, Vigo, Spain
15 Gastroenterology, Hospital Alvaro Cunqueiro, Vigo, Spain
16 Instituto de Investigacion e Innovacion en Bioingeniera, I3B, Universitat Politecnica de Valencia, Valencia, Valencian Community, Spain

## Statements and Additional Information

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Contributors' Statement** A.N.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing—original draft, Writing—review & editing. R.N.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing—original draft. N.A.-L.: Data curation, Resources, Validation, Writing—original draft. C.C.: Conceptualization, Data curation, Methodology, Resources, Validation, Writing—original draft. C.P.F.: Data curation, Resources, Validation. B.G.S.: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Validation, Writing—original draft. A.G.: Data curation, Resources. V.-A.J.-G.: Data curation, Resources. L.L.G.: Data curation, Resources, Validation. Á.C.-Á.: Data curation, Resources, Validation. J.V.M.: Data curation, Resources, Validation. A.G.V.: Data curation, Resources, Validation. M.L.-S.: Data curation, Resources, Validation, Writing—original draft. L.R.S.: Data curation, Resources, Validation. F.S.C.: Data curation, Resources, Validation. J.L.M.H.: Data curation, Resources, Validation. J.G.L.: Data curation, Resources, Validation. O.N.R.: Data curation, Resources, Validation. J.E.V.: Data curation, Resources, Validation. M.J.M.L.: Data curation, Resources, Validation. A.M.Z.: Data curation, Resources, Validation. A.M.-T.: Data curation, Software, Visualization. M.B.-B.: Data curation, Formal analysis, Supervision. V.N.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Writing—original draft, Writing—review & editing. V.P.B.: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing.

## References

1 Rokkas T, Papaxoinis K, Triantafyllou K, Pistiolas D, Ladas SD. Does purgative preparation influence the diagnostic yield of small bowel video capsule endoscopy?: A meta-analysis. Am J Gastroenterol 2009;104(01):219–227. 10.1038/ajg.2008.63

2 Domagk D, Despott E, Spada C, et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) technical review. Endoscopy 2018. 10.1055/a-0576-0566

3 Enns RA, Hookey L, Armstrong D, et al. Clinical practice guidelines for the use of video capsule;endoscopy. Gastroenterology 2017;152(03):497–514. 10.1053/j.gastro.2016.12.032

4 Dray X, Houist G, Le Mouel JP, et al. Prospective evaluation of third-generation small bowel capsule endoscopy videos by independent readers demonstrates poor reproducibility of cleanliness classifications. Clin Res Hepatol Gastroenterol 2021;45(06). 10.1016/j.clinre.2020.101612

5 Leenhardt R, Koulaouzidis A, Histace A, et al. Key research questions for implementation of artificial intelligence in capsule endoscopy. Ther Adv Gastroenterol 2022:15. 10.1177/17562848221132683

6 Noorda R, Nevárez A, Colomer A, Pons Beltrán V, Naranjo V. Automatic evaluation of degree of cleanliness in capsule endoscopy based on a novel CNN architecture. Sci Rep 2020;10(01):1–13. 10.1038/s41598-020-74668-8

7 Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019:95. 10.1016/j.jbi.2019.103208

8 Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(02):377–381. 10.1016/j.jbi.2008.08.010

9 Pons Beltrán V, González Suárez B, González Asanza C, et al. Evaluation of different bowel preparations for small bowel capsule endoscopy: a prospective, randomized, controlled study. Dig Dis Sci 2011;56(10):2900–2905. 10.1007/s10620-011-1693-z

10 Pennazio M, Rondonotti E, Despott EJ, et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) guideline – update 2022. Endoscopy 2022. 10.1055/a-1973-3796

11 Yung DE, Rondonotti E, Sykes C, Pennazio M, Plevris JN, Koulaouzidis A. Systematic review and meta-analysis: is bowel preparation still necessary in small bowel capsule endoscopy? Expert Rev Gastroenterol Hepatol 2017;11(10):979–993. 10.1080/17474124.2017.1359540

12 Gkolfakis P, Tziatzios G, Dimitriadis GD, Triantafyllou K. Meta-analysis of randomized controlled trials challenging the usefulness of purgative preparation before small-bowel video capsule endoscopy. Endoscopy 2018;50(07):671–683. 10.1055/s-0043-125207

13 Marmo C, Riccioni ME, Pennazio M, Antonelli G, Spada C, Costamagna G. Small bowel cleansing for capsule endoscopy, systematic review and meta- analysis: timing is the real issue. Dig Liver Dis 2022. 10.1016/j.dld.2022.07.002

14 Ponte A, Pinho R, Rodrigues A, Carvalho J. Review of small-bowel cleansing scales in capsule endoscopy: a panoply of choices. World J Gastrointest Endosc 2016;8(17):600. 10.4253/wjge.v8.i17.600

15 Cortegoso Valdivia P, Deding U, Bjørsum-Meyer T, et al. Inter/intra-observer agreement in video-capsule endoscopy: are we getting it all wrong? A systematic review and meta-analysis. Diagnostics 2022;12(10). 10.3390/diagnostics12102400

16 Van Weyenberg SJB, De Leest HTJI, Mulder CJJ. Description of a novel grading system to assess the quality of bowel preparation in video capsule endoscopy. Endoscopy 2011;43(05):406–411. 10.1055/s-0030-1256228

17 Leenhardt R, Souchaud M, Houist G, et al. A neural network-based algorithm for assessing the cleanliness of small bowel during capsule endoscopy. Endoscopy 2021. 10.1055/a-1301-3841

18 Nam JH, Oh DJ, Lee S, Song HJ, Lim YJ. Development and verification of a deep learning algorithm to evaluate small-bowel preparation quality. Diagnostics 2021;11(06). 10.3390/diagnostics11061127

19 Houdeville C, Leenhardt R, Souchaud M, et al. Evaluation by a machine learning system of two preparations for small bowel capsule endoscopy: the BUBS (burst unpleasant bubbles with Simethicone) study. J Clin Med 2022;11(10):2822. 10.3390/jcm11102822

20 Rahmi G, Cholet F, Gaudric M, et al. Effect of different modalities of purgative preparation on the diagnostic yield of small bowel capsule for the exploration of suspected small bowel bleeding: a Multicenter randomized controlled trial. Am J Gastroenterol 2022;117(02):327–335. 10.14309/ajg.0000000000001597

21 Io Lei I, Cortegoso Valdivia P, Gualandi N, et al. Identifying the optimal bowel preparation for small-bowel capsule endoscopy: a meta-review and umbrella meta-analysis. Gastrointest Endosc 2026;103(04):672–686.e24. 10.1016/j.gie.2025.11.005