Validation and Application of Predictive Models for Inadequate Bowel Preparation in Colonoscopies in a Tertiary Hospital Population

Edgar Afecto    Ana Ponte    Sónia Fernandes    Catarina Gomes
João Paulo Correia    João Carvalho

Department of Gastroenterology, Centro Hospitalar Vila Nova de Gaia/Espinho, Vila Nova de Gaia, Portugal

Keywords
Colonoscopy · Bowel preparation · Colonoscopy quality criteria

Abstract
Background: Bowel preparation is a major quality criterion for colonoscopies. Models developed to identify patients with inadequate preparation have not been validated in external cohorts. We aim to validate these models and determine their applicability. Methods: Colonoscopies between April and November 2019 were retrospectively included. Boston Bowel Preparation Scale $\geq 2$ per segment was considered adequate. Insufficient data, incomplete colonoscopies, and total colectomies were excluded. Two models were tested: model 1 (tricyclic antidepressants, opioids, diabetes, constipation, abdominal surgery, previous inadequate preparation, inpatient status, and American Society of Anesthesiology [ASA] score $\geq 3$); model 2 (co-morbidities, tricyclic antidepressants, constipation, and abdominal surgery). Results: We included 514 patients (63% males; age 61.7 ± 15.6 years), 441 with adequate preparation. The main indications were inflammatory bowel disease (26.1%) and endoscopic treatment (24.9%). Previous surgery (36.2%) and ASA score $\geq 3$ were the only identified predictors for inadequate preparation in this study ($p < 0.001$, OR 3.28). The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of model 1 were 60.3, 64.2, 21.8, and 90.7%, respectively. Model 2 had a sensitivity, specificity, PPV, and NPV of 57.5, 67.4, 22.6, and 90.5%, respectively. The AUC for the ROC curves was 0.62 for model 1, 0.62 for model 2, and 0.65 for the ASA score. Conclusions: Although both models accurately predict adequate bowel preparation, they are still unreliable in predicting inadequate preparation and, as such, new models, or further optimization of current ones, are needed. Utilizing the ASA score might be an appropriate approximation of the risk for inadequate bowel preparation in tertiary hospital populations.

© 2021 Sociedade Portuguesa de Gastroenterologia. Published by S. Karger AG, Basel

Validação e aplicação de modelos preditivos para preparação intestinal inadequada numa população de um hospital terciário

Palavras Chave
Colonoscopia · Preparação intestinal · Critérios de qualidade na colonoscopia
Two recent systematic reviews and meta-analyses [2, 3] attempted to identify predictive factors for inadequate bowel preparation. Patient age, male gender, medical history (chronic constipation, hypertension, diabetes, cirrhosis, stroke, and dementia), and current medication (opiates and tricyclic antidepressants) were identified as risk factors, while previous abdominal surgery, previous inadequate bowel preparation, and body mass index were not consistent predictors.

Based on these findings, three predictive models [4–6] were developed with the aim of identifying patients at risk for inadequate bowel preparation. To the best of our knowledge, none of these models have yet been applied outside of their development/validation cohorts and in clinical practice.

Therefore, this study aims to validate these inadequate bowel preparation predictive models in our population and to determine their applicability in clinical practice.

Materials and Methods

Study Population and Data Collection

This study was conducted in a tertiary hospital in northern Portugal. For the purpose of this retrospective cohort study, the authors considered all patients who underwent an elective colonoscopy between April and November 2019. Patient data was collected through database search and clinical records, concerning sex, age, indication for colonoscopy, type of preparation used, American Society of Anesthesiology (ASA) score, simple medical history (diabetes, cirrhosis, neurological disorders, abdominal or pelvic surgery), medication history (opioids, tricyclic antidepressants), previous inadequate bowel preparation, chronic constipation history (defined as fewer than three bowel movements per week), and current hospitalization at the time of colonoscopy.

Bowel preparation was considered adequate if every colon segment scored at least 2 points in the Boston Bowel Preparation Scale (BBPS). The exclusion criteria were absence of BBPS in the exam report, incomplete colonoscopy (for a reason other than inadequate bowel preparation), total colectomy, and incomplete patient data.

Bowel Preparation and Other Interventions

According to the implemented protocol in our hospital, all patients were provided with written and oral instructions regarding bowel preparation. A low-fiber diet was started in the 2 days prior to the procedure. Per protocol, patients are allowed to choose between low-volume polyethylene glycol (PEG) + ascorbic acid, high-volume PEG or sodium picosulfate (healthy patients, without comorbidities). A split-dose bowel preparation regimen was used for colonoscopies scheduled for the morning period, while a same-day regimen was used for afternoon colonoscopies. Additionally, a nurse was available for a face-to-face consultation with every patient that had doubts or required further instructions regarding bowel preparation regimens.
Two predictive models, previously published, were tested in our population. Model 1 by Dik et al. [5] considers tricyclic antidepressants or opioids use, diabetes, constipation, previous abdominal surgery, previous inadequate preparation, inpatient status, and ASA score ≥3. Model 2 by Gimeno-García et al. [6] considers comorbidities (diabetes, cirrhosis, and stroke history), tricyclic antidepressants use, constipation, and previous abdominal surgery. A third model, developed by Hassan et al. [4], was not tested as it did not use a validated bowel preparation scale.

**Bowel Preparation Predictor Models**

**Statistical Analysis**

Continuous variables are expressed as medians and interquartile ranges (IQR), while categorical variables are expressed as frequencies and percentages. All potentially inadequate bowel preparation factors were subjected to univariate ($\chi^2$ for categorical and $t$ test for continuous variables) and multivariate analysis (logistic regression). Two-tailed $p$ values were considered statistically significant if $<0.05$. A subset analysis comparing high- and low-volume PEG was performed regarding ASA score, age, comorbidities, constipation, previous surgery, current inpatient status, and previous colonoscopy with inadequate preparation as we postulated.
that old patients with more comorbidities were probably being offered more high-volume PEG preparations due to safety concerns.

The discriminative power of both models in predicting inadequate bowel preparation was determined by calculating the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the area under the curve (AUC) of the receiver operating characteristic (ROC) curves.

The $\chi^2$ test of goodness-of-fit was performed to evaluate whether the sample size was adequate, using the G*Power software and data available in the literature [5, 6]. This analysis revealed that a power of 95% for model 1 would have been achieved with 169 patients, while for model 2, enrolling 423 patients would allow for a power of 90%; in the present study, 514 individuals were included, allowing adequate power to test both models.

All statistical analyses were performed in IBM SPSS Statistics v22.0 (SPSS Inc., Chicago, IL, USA) and G*power v3.1.9.7 (Heinrich-Heine-Universität Düsseldorf, Germany).

Results

We included 514 patients, 324 (63.0%) of which were males. The median age was 64 years (IQR 53–73.25). The most common indications for colonoscopy were inflammatory bowel disease ($n = 134$, 26.1%), endoscopic treatment ($n = 128$, 24.9%), and cancer/polypectomy follow-up ($n = 125$, 24.4%). Previous abdominal/pelvic surgery was the most common comorbidity ($n = 186$, 36.2%), followed by diabetes ($n = 110$, 21.4%). The majority ($n = 371$, 72.2%) of patients underwent bowel cleansing with PEG-based solution. A more extensive description of patient clinical data and baseline characteristics can be found in Table 1.

Adequate bowel preparation, as defined in the Methods section, was observed in 441 (85.8%) patients. The median total BBPS score was 7 points (IQR 6–9), and the median right colon, transverse colon, and left colon segment scores were 2, 3, and 2 points, respectively. This data is summarized in Table 2.

Table 2. Bowel preparation

<table>
<thead>
<tr>
<th>Bowel preparation, n (%)</th>
<th>441 (85.8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequate</td>
<td>441 (85.8)</td>
</tr>
<tr>
<td>Inadequate</td>
<td>73 (14.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total BBPS, mean ± SD</th>
<th>7.04±1.85</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–2, n (%)</td>
<td>10 (1.9)</td>
</tr>
<tr>
<td>3–5 and 6 (at least 1 segment &lt;2), n (%)</td>
<td>63 (12.3)</td>
</tr>
<tr>
<td>6–9, n (%)</td>
<td>441 (85.8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BBPS per segment, mean ± SD</th>
<th>2.28±0.72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right colon</td>
<td>2.44±0.67</td>
</tr>
<tr>
<td>Transverse colon</td>
<td>2.32±0.67</td>
</tr>
<tr>
<td>Left colon</td>
<td>2.32±0.67</td>
</tr>
</tbody>
</table>

On bivariate analysis, we found an association between inadequate preparation and opioid use ($p = 0.027$), diabetes mellitus ($p = 0.001$), chronic constipation ($p = 0.011$), previous inadequate preparation ($p = 0.034$), and ASA score ($p < 0.001$). On further multivariate analysis, the same effect was observed only for those with higher ASA scores ($p < 0.001$, OR 3.28, 95% CI 2.04–5.28). No association was found regarding age, sex, volume of PEG used (high versus low volume), tricyclic antidepressant use, cirrhosis, neurologic comorbidities, previous surgery (even when separated by colon resection vs. other intra-abdominal surgeries), and inadequate bowel preparation. This information is also displayed in Table 1.

Model 1 predicted a total of 202 patients as having inadequate preparation, 44 of which were correctly predicted as such. On the other hand, 312 were predicted as adequate preparation, 283 of which were correctly predicted. The sensitivity, specificity, PPV, and NPV of model 1 were 60.3, 64.2, 21.8, and 90.7%, respectively. The AUC for the ROC curve of this model was 0.62 (95% CI 0.55–0.69).

Model 2 predicted 186 patients as having inadequate bowel preparation, 42 of which were correctly predicted as such. On the other hand, 328 were predicted as adequate preparation, 297 of which were correctly predicted. Model 2 had a sensitivity, specificity, PPV, and NPV of 57.5, 67.4, 22.6, and 90.6%, respectively. The AUC for the ROC curve of this model was 0.62 (95% CI 0.55–0.69).
model was 0.62 (95% CI 0.55–0.70). ROC curves for both models are shown in Figure 1.

A subset analysis comparing high- and low-volume PEG regimens showed that higher ASA scores \((p = 0.001)\) and inpatient status \((p = 0.008)\) were significantly associated (on bivariate and multivariate analysis) with the use of higher-volume PEG regimens. Higher age was also associated with higher-volume PEG on bivariate analysis, but this was not confirmed on multivariate analysis.

Lastly, as the ASA score was the only predictive factor in our study, we tested its accuracy in predicting inadequate preparation. Utilizing an ASA score >2, 124 patients were predicted as having inadequate preparation, 34 of which were correctly predicted. On the other hand, 390 were predicted as having adequate preparation, 351 of which were correctly predicted. This translates as a sensitivity, specificity, PPV, and NPV of 46.6, 79.6, 27.4, and 90.0%, respectively. The AUC for the ASA score was 0.65 \((p < 0.001, 95\% \text{ CI} 0.58–0.72)\). This ROC curve is also presented in Figure 1.

### Discussion

In this study, we were able to replicate the data published by the authors of these scores, with similar, although slightly worse, results [5, 6].

The ongoing search for factors that can influence bowel preparation and accurately identify these patients resulted in these two previously published scores. Although the results seemed promising, some limitations are easily identified and were further confirmed in our study. More than half of the patients with inadequate preparation were identified, but there is still a significant group of patients that were not identified by either score (39.7% for model 1 and 42.5% for model 2) and in which we thus could not intervene. Although the argument can be made that all patients identified by these scores as having inadequate preparation could be offered more intensive regimens, it is also true that we would be subjecting a non-negligible number of patients to unnecessary interventions (which could further reduce compliance), as is demonstrated by the observed very low PPVs and high false positives for both scores. As such, the scores demonstrated a low value in predicting inadequate bowel preparation.

On the other hand, both scores were found to have a very high NPV, which means they could be useful in determining which patients most likely do not require additional interventions.

When comparing the original model 1 study, by Dik et al. [5], some methodological similarities and differences can be pointed out. While most of our exams were performed in an inflammatory bowel disease or endoscopic treatment setting, the original study included mostly patients undergoing screening or symptom investigation, which may induce a difference in adhesion to the bowel cleansing protocols due to different populations. Additionally, in this study, BBPS was defined as inadequate if total <6, and no reference was made to segments scoring 1 point with a total of 6, which we considered as inadequate preparation in our paper. Although the majority of patients in both our study and the original study are ASA class 1 or 2, while Dik et al. [5] only had 4.5% of patients with ASA ≥3, we had 24.1% scoring ≥2 points, which can be explained by an overall increased prevalence of comorbidities in our population. Overall, while there are differences in the populations being compared, model 1 performed worse in our study, with lower sensitivity (0.60 vs. 0.66), specificity (0.79 vs. 0.64), PPV (0.22 vs. 0.29), and NPV (0.91 vs. 0.95).

Regarding the original study for model 2, by Gimeno-García et al. [6], similar limitations can be described. Most examinations were also performed in a screening/symptom investigation setting, with equivalent conclusions regarding applicability. In this study, the median age of patients (60 vs. 64 in our study) was lower and comorbidities were present in a lower proportion (21.8 vs. 24.1%). Patients with dementia and previous history of inadequate bowel preparation were excluded in this study but included in our study because we believe these patients are at greater risk of inadequate bowel preparation. Additionally, the proportion of opioid (4.8 vs. 1.8%) or tricyclic antidepressant (8.2 vs. 1.4%) use was significantly higher in their population. On the other hand, as was the case with the model 1 study [5], the study by Gimeno-García et al. [6] was prospective in nature, contrasting with our retrospective study and its inherent biases. With these differences summarized, differences in model 2 accuracy were expected and were observed as a higher sensitivity (0.58 vs. 0.50) and NPV (0.91 vs. 0.88) but a lower specificity (0.67 vs. 0.80) and PPV (0.23 vs. 0.36).

Conversely, chronic constipation and abdominal surgery were not identified as predictors. This could be explained by the retrospective nature of our study, mostly in the case of chronic constipation, as we could not always use objective definitions for these categories due to incomplete data. In concordance with both studies [5, 6], we considered both partial colectomies and other intra-abdominal/pelvic surgeries the same for the purpose of
this study, and, as such, different relationships between surgeries and inadequate preparation can arise (as we do not know what surgeries the patients in the original studies had in order to make a comparison). Additionally, a sub-analysis regarding type of surgery (colon/rectum resection or urogenital surgery) found no relationship with inadequate preparation. Nevertheless, the relationship between previous surgery and inadequate preparation is controversial, as two meta-analyses with a significant number of patients failed to demonstrate any relationship [2, 3], probably due to the heterogeneity of definitions used and patients included.

The models previously described are probably more useful outside of tertiary specialized hospitals, as the population in our study (older, more comorbidities, and lower proportion of screening/symptom investigation) demonstrated substantially different results and lower AUC values than previously published.

In terms of previously identified inadequate bowel preparation predictors, we were only able to identify higher ASA scores as a predictive factor for an inadequate preparation in our population. By utilizing the ASA score alone, we demonstrated a predictive power similar but slightly better to the two models tested (AUC 0.65 vs. 0.62 on both models tested). The ASA score is widely used and easy to apply in clinical practice and categorizes patients according to their comorbidities. In populations such as ours (a tertiary hospital), utilizing the ASA score in order to triage patients who should be paid more attention regarding bowel preparation is an easier and quicker method than the two scores analyzed in this paper.

When aiming to optimize bowel preparation, several steps should be taken by all patients. Patients should adopt a clear-liquid or low-fiber diet on the previous day as both are equally effective [7, 8], although a low-fiber diet is associated with higher tolerability and willingness to repeat the exam. Bowel preparation should be undertaken as a split-dose regimen for next-day procedures or a same-day regimen for afternoon procedures [1, 9, 10]. The bowel preparation utilized (high-volume PEG, low-volume PEG, or non-PEG regimen if clinically validated) can be chosen according to patient preference, as there seems to be no difference in efficacy between regimens [11]. High-volume regimens offer better safety profiles with the trade-off of diminished tolerability (and thus more inadequate preparations), which might be more relevant in older patients with more comorbidities [12].

With all these previous measures applied, bowel preparation is more likely to be optimized, even in patients who are thought to be at risk (such as ASA >2). Nevertheless, additional measures, such as enhanced bowel preparation instructions, should be applied—such as a face-to-face or telephone nursing consultation [13, 14]. Although it may seem reasonable to prescribe additional laxatives or high-volume preparations in constipated individuals, the current available evidence shows no difference between regimens in these patients [15]. In case of a previous inadequate preparation, a modifiable reason for the failure of the chosen regimen should be sought before prescribing a different regimen or additional measures, such as nausea/vomiting or poor adherence due to patient- or preparation-related factors.

Several limitations can be readily identified, mostly due to the retrospective nature of our study and its inherent biases, such as the likelihood of suffering from missing data. Patient compliance or tolerance to bowel preparation was not registered, but nevertheless, our reported proportion of adequate bowel preparation was similar to previously published literature [5, 6] and nearly achieved the ≥90% recommended threshold [16], which probably indicates an adequate (but not perfect) compliance. Additionally, not all preparation regimens were registered, as this was not practice in our hospital at this time period, but it is reasonable to assume that most of these patients underwent a PEG solution, although the proportion of which cannot be inferred due to missing data. Regarding volume of preparation, the utilization of higher-volume PEG solutions was higher in older patients (although not significantly so), inpatients, and higher ASA scores: all inpatients are prescribed higher-volume PEG, as it is the only solution available in our hospital; for higher ASA and age, we postulate that due to comorbidities and advanced age, these patients were probably recommended by the nursing staff or the community pharmacist to undergo higher-volume preparations for safety reasons. Although this may introduce a bias, we believe it not to be significant since the proportion of adequate preparation was the same between groups (high vs. low volume). Lastly, in our population, 26.1 and 24.9% of colonoscopies were performed in inflammatory bowel disease and endoscopic therapy settings, which limits the generalization of our data and indicates that these scores are not suitable for a tertiary hospital population such as ours.

In conclusion, although both models are capable of predicting more than half of the patients with inadequate bowel preparation, they are unable to do so in a reliable manner and, therefore, almost half of those requiring more intensive regimens are not identified when using these models. Further improvement of these models or development of new ones is necessary before they can be
applied in clinical practice. Utilizing the ASA score might be an appropriate approximation of the risk for inadequate bowel preparation in tertiary hospital populations.

Statement of Ethics

The study protocol was approved by the local ethics committee. The research was conducted ethically in accordance with the Declaration of Helsinki 2014 [9].

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

References


Funding Sources

This work was not funded.

Author Contributions