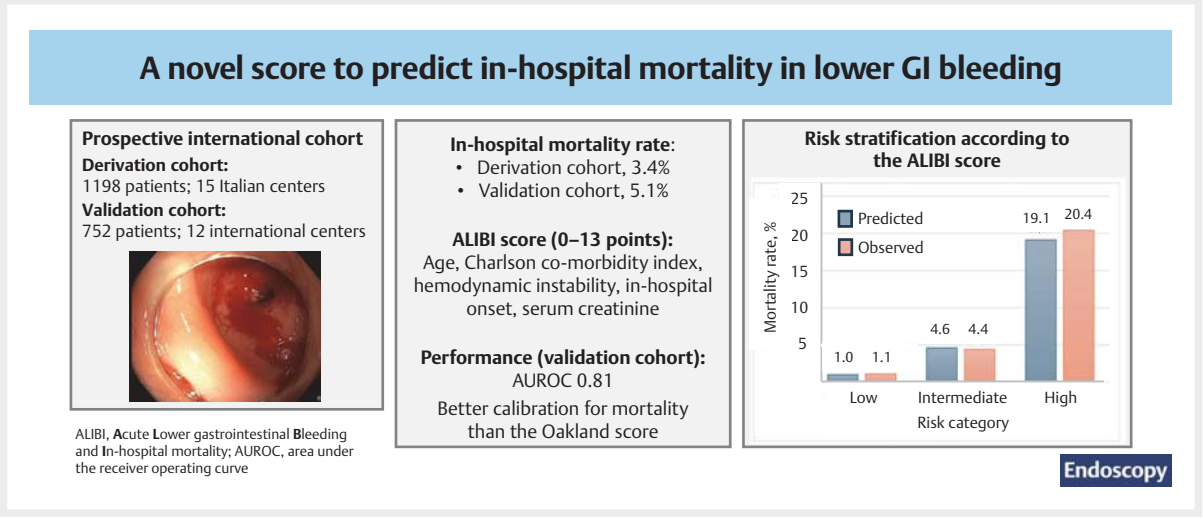
Accepted Manuscript online: 2025-02-17 Article published online: 2025-04-04

In-hospital mortality in patients with lower gastrointestinal bleeding: development and validation of a prediction score

GRAPHICAL ABSTRACT



Authors

Elton Dajti[1,2], Leonardo Frazzoni[3], Sílvia Castellet-Farrús[4], Jordi Guardiola[4], Emanuele Sinagra[5], Andrea Anderloni[6], Francesco Ferrara[7], Paraskevas Gkolfakis[8], Marine Camus Duboc[9], Francesco Vito Mandarino[10], Anahita Sadeghi[11], Vicente Lorenzo-Zúñiga[12], Sandra Perez[13], Konstantinos Triantafyllou[14], Maria Paula Curado[15], Antonio Facciorusso[16], Giulia Collatuzzo[2], Cesare Hassan[17,18], Franco Radaelli[19], Lorenzo Fuccio[12], on behalf of the ALIBI Study Group

The ALIBI Study Group

Aurelio Mauro, Alberto Busatto, Georgios Tziatzios, Antonia Panagaki, Dario Raimondi, Francesca Rossi, Giuseppe Conoscenti, Rita Alloro, Silvio Danese, Reza Malekzadeh, Elena Coello, Enrique Rodriguez de Santiago, Bruna Elisa Catin Kupper, Samuel Aguiar, Wilson Toshihiko Nakagawa, Paolo Boffetta, Giovanni Barbara, Cristina Mosconi, Alessandro Repici, Alessandro Mussetto, Cristiano Spada, Gianpiero Manes, Sergio Segato, Alessandro Musso, Emilio Di Giulio, Mauro Manno, Germana De Nucci, Virginia Festa, Alfredo Di Leo, Mario Marini, Luca Ferraris

Institutions

1 Gastroenterology Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
2 Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Bologna, Italy
3 Gastroenterology and Digestive Endoscopy Unit, Forlì-Cesena Hospitals, AUSL Romagna, Forlì, Italy
4 Department of Gastroenterology, Hospital Bellvitge, L'Hospitalet, Barcelona, Spain
5 Gastroenterology and Endoscopy Unit, Fondazione Istituto S. Raffaele-G. Giglio, Cefalù, Italy
6 Gastroenterology and Digestive Endoscopy Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
7 Gastroenterology Unit, University Hospital of Padova, Padova, Italy
8 Department of Gastroenterology, "Konstantopoulio-Patision" General Hospital of Nea Ionia, Athens, Greece
9 Centre for Digestive Endoscopy, Sorbonne University, Saint-Antoine Hospital, APHP, Paris, France
10 Department of Gastroenterology and Endoscopy, IRCCS San Raffaele Hospital, Milan, Italy
11 Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran
12 Endoscopy Unit, Hospital Universitari i Politècnic La Fe/IISLaFe, Valencia, Spain
13 Gastroenterology and Hepatology, Hospital Universitario Ramón y Cajal, Universidad de Alcalá, IRYCIS, Madrid, Spain
14 Hepatogastroenterology Unit, Attikon University General Hospital, Athens, Greece

15 Colorectal Cancer Department, AC Camargo Cancer Center, São Paulo, Brazil
16 Gastroenterology, Department of Medical and Surgical Sciences, University of Foggia, Foggia, Italy.
17 Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Ital
18 IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy
19 Gastroenterology Unit, Valduce Hospital, Como, Italy

submitted 23.10.2024
accepted after revision 17.2.2025
accepted manuscript online 17.2.2025
published online 4.4.2025

Bibliography
Endoscopy 2025; 57: 839–848
DOI 10.1055/a-2541-2312
ISSN 0013-726X

Georg Thieme Verlag KG, Oswald-Hesse-Straße 50, 70469 Stuttgart, Germany

Supplementary Material
Supplementary Material is available at https://doi.org/10.1055/a-2541-2312

Correspondig author
Lorenzo Fuccio, MD, Gastroenterology Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
lorenzo.fuccio3@unibo.it

ABSTRACT

Background Lower gastrointestinal bleeding (LGIB) is a common condition linked to increased morbidity, healthcare costs, and mortality. Currently, no prospectively validated prognostic model exists to predict mortality in patients with LGIB. Our aim was to develop and validate a risk score that could accurately predict in-hospital mortality of patients admitted for LGIB.

Methods Patient data from a nationwide cohort study in 15 centers in Italy (2019–2020) were used to derive the risk score, the Acute Lower gastrointestinal Bleeding and In-hospital mortality (ALIBI) score; the model was then externally validated in a cohort of consecutive patients hospitalized for LGIB in 12 centers from six countries (Italy, Spain, France, Greece, Iran, and Brazil) from 2022 to 2024. The main outcome was in-hospital mortality; we also reported rebleeding rates and the in-hospital mortality rate stratified by risk score and timing of colonoscopy.

Results Among 1198 patients in the derivation cohort, 105 (8.8%) re-bled and 41 (3.4%) died. Age, Charlson Comorbidity Index, in-hospital onset, hemodynamic instability, and creatinine level were independent predictors of in-hospital mortality. The model demonstrated excellent discrimination (area under the receiver operating curve [AUROC] 0.81, 95%CI 0.75–0.87) and calibration. In the validation cohort (n = 752 patients), the model's good discrimination (AUROC 0.79, 95%CI 0.72–0.86) and calibration were confirmed. Patients were categorized as low (0–4 points; 1% mortality), intermediate (5–9 points; 4.6% mortality), or high risk (10–13 points; 19.1% mortality).

Conclusion A new validated score effectively predicts in-hospital mortality in patients with LGIB, aiding in their risk stratification and management.

Introduction

Acute lower gastrointestinal bleeding (LGIB) is a common cause of hospitalization and is associated with high morbidity, healthcare burden, and cost [1]. It accounts for up to 20 % of all admissions for gastrointestinal bleeding, but its incidence and mortality may have risen in recent decades owing to increases in the average age of the population and the use of antithrombotic agents [2, 3]. The severity of LGIB varies widely from self-limiting hematochezia to major bleeding with anemia and hemodynamic instability; older age, co-morbidities, smoking, alcohol consumption, and the use of nonsteroidal anti-inflammatory drugs (NSAIDs) or aspirin are the main risk factors for bleeding [4]. Mortality from LGIB is however a relatively rare event, ranging from 3.4% to 8.8% [5, 6, 7]. Therefore, risk assessment at initial clinical presentation is of paramount importance in predicting unfavorable outcomes, the need for hospitalization and interventions, and death.

Unlike for upper gastrointestinal bleeding, clinical risk scores developed specially to predict morbidity and mortality in patients with LGIB are limited [8]. The Oakland score is the most widely used score in this setting; it was designed to identify patients with LGIB who could safely be discharged for outpatient investigations [4, 5, 9]. Other scores have been developed to predict severe bleeding or composite outcomes [4, 8, 10], but the definition of the outcome is heterogeneous, some variables are subjective or unavailable at the emergency setting (e. g. albumin), and prospective validation is lacking. To date, no risk score has been developed to predict a hard clinical outcome such as in-hospital mortality in patients with LGIB, and the discriminative ability of the above scores is insufficient to evaluate this outcome [5, 11].

We aimed to develop and validate a prognostic score that could accurately predict the risk of mortality in hospitalized patients with LGIB. We also described the mortality rates of patients with LGIB stratified by their risk score at baseline and the timing of colonoscopy.

Methods

Study design and population

To develop the new prediction model, we used data from consecutive patients with LGIB hospitalized across 15 Italian centers between April 2019 and March 2020. As previously reported [12], this was a prospective observational cohort study aiming to benchmark the etiology, management, and outcomes of patients with LGIB in Italy. Briefly, the main inclusion criteria were: (i) age ≥ 18 years; (ii) recent (< 3 days) onset of LGIB, either leading to emergency room attendance with subsequent admission or occurring in patients already hospitalized for other reasons. Patients presenting with melena who did not undergo upper gastrointestinal endoscopy or who had a proven or probable source of bleeding identified during the upper gastrointestinal endoscopy were excluded from the analysis.

To externally validate the new model both temporally and geographically, we specifically conducted a prospective international cohort study across 12 centers and six countries (Italy, Spain, France, Greece, Brazil, and Iran) from December 2022 until February 2024. None of the Italian centers participating in the external validation cohort had participated in the derivation cohort. The eligibility criteria, the clinical setting, and the outcome measures were the same as those used in the derivation cohort. The number of patients enrolled in each participating center is reported in **Table 1 s**, see online-only Supplementary material.

Data collection

For each patient, we collected the following data at the time of the LGIB presentation: age, sex, co-morbidities, symptoms, hemodynamic status, setting (emergency room vs. hospitalized, weekday vs. weekend diagnosis), laboratory tests (hemoglobin levels, platelet count, coagulation, serum creatinine), list of medications, history of a previous admission for LGIB. Hemodynamic instability was defined as systolic blood pressure < 100 mmHg and/or heart rate > 100 beats per minute at presentation. The variable "in-hospital onset" was considered to be positive if the bleeding episode occurred in a patient who had already been hospitalized for another medical reason, as opposed to a patient who presented to the emergency room and was hospitalized specifically for the episode of LGIB. Data on co-morbidities were defined using the Charlson Co-morbidity Index (CCI) with no additional points assigned according to age, as reported in **Table 2 s**.

In the validation cohort, we also collected the variables required to calculate the Oakland score that were not available in the derivation cohort (blood pressure, heart rate, rectal examination findings). We also collected data on patient management (first examination performed, type and timing of endoscopy, need for endoscopic, radiologic, or surgical treatment, management of anticoagulation or antiplatelet drugs, need for transfusion), cause of bleeding, and patient outcomes (rebleeding and mortality). Rebleeding was defined as one of the following: overt LGIB, new hemodynamic instability or hemoglobin level drop ≥ 2 g/dL after initial stabilization. ▶ **Fig. 1** shows example images from two patients with LGIB who were managed by interventional radiology and endoscopically, respectively.

Statistical analysis

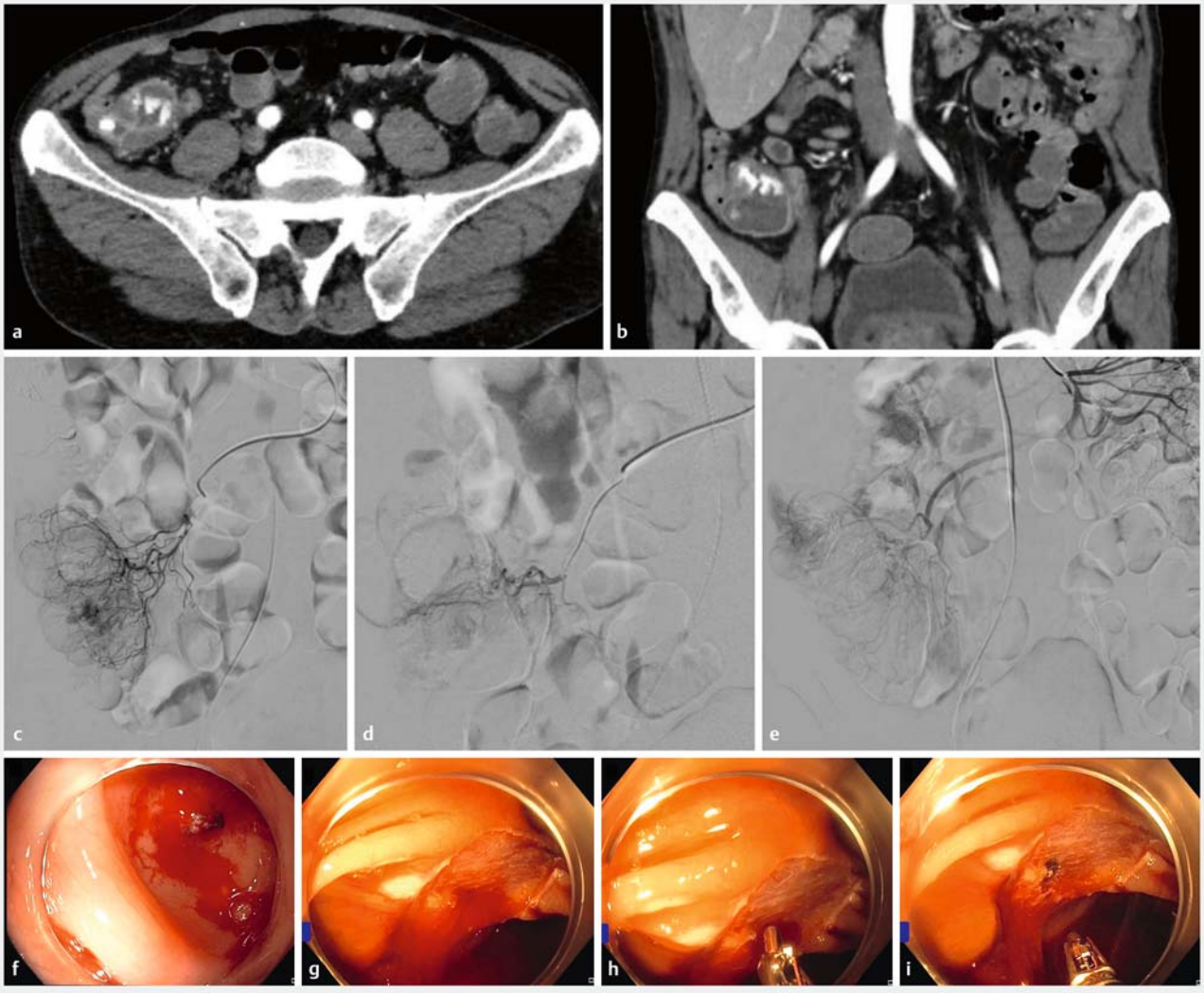
Categorical data were expressed as numbers (percentages), and continuous variables as medians (interquartile range [IQR]). The statistical plan for model development and validation is reported in **Appendix 1 s**.

The main outcome was in-hospital mortality. In the derivation cohort, logistic regression analysis was performed starting from a prespecified set of fifteen candidate variables that were both readily available and plausibly related to the dependent variable (in-hospital mortality): age, sex, CCI, previous hospitalization for LGIB, antiplatelet therapy, anticoagulation therapy, use of NSAIDs, LGIB onset setting (out-of-hospital vs. in-hospital), weekend diagnosis, symptoms at presentation, hemodynamic instability, hemoglobin levels, platelet count, international normalized ratio, and serum creatinine. After evaluation of multicollinearity, multivariable logistic regression analyses were performed on variables that reached a significance of $P < 0.1$ on univariable analysis.

The final multivariable model was built from the set of candidate variables by removing the predictors based on their *P* values, in a stepwise manner. The linearity assumption underlying the logistic regression model was tested using the Box–Tidwell model [13]. Model discrimination was assessed by the area under the receiver operating characteristic (AUROC) curve; the DeLong test was performed to test the equality between two or more AUROC curves [14]. Calibration was assessed by a calibration curve showing the relationship between the estimated risk (on the *x* axis) and the observed proportion of events (*y* axis). The accuracy of the final model was internally validated using 1000 bootstrap samples [15] and the optimism-corrected AUROC was calculated [16, 17]. We used the regression coefficient estimates from the multivariable model to derive the equation for predicting the main outcome, which was then validated in the external validation cohort.

Finally, we constructed a new scoring system following the methodology proposed by Sullivan et al. in the Framingham Study [18] to enhance the use of the model in routine clinical practice (**Appendix 1 s**). The score's discriminative ability and calibration for predicting the main outcome was evaluated as mentioned above and externally validated in an independent cohort. Patients were then stratified into levels of low, intermediate, and high risk, with thresholds reflecting clinically meaningful (at least two-fold) gradients in risk from one group to the next.

In the validation cohort, we compared the performance of our score with that of the Oakland score. Moreover, we reported the rate of in-hospital mortality according to the risk categories (low vs. intermediate-to-high), colonoscopy procedure (yes vs. no), and timing (early < 24 hours vs. delayed). The rate of missing data in the validation set was very low (0.8 %), so we excluded these patients from the analysis and did not use any imputation method.



▶ **Fig. 1** Images from two patients with typical lower gastrointestinal bleeding. **a–e** Patient #1, who had severe lower gastrointestinal bleeding. **a,b** Computed tomography images showing arterial bleeding in the cecum; **c–e** angiographic images showing: **c** confirmation of the arterial bleeding; **d** super-selective glue embolization being performed; **e** resolution of the bleeding. **f–i** Endoscopic images from patient #2, who had bleeding due to endoscopic submucosal dissection showing: **f,g** the site of bleeding; **h** combined mechanical and energy-based hemostasis being applied; **i** subsequent control of the bleeding.

All *P* values refer to two-tailed significance tests. $P < 0.05$ was considered significant. Statistical analyses were performed with Stata/SE, version 18 for Windows (Stata Corp., Texas, USA). We followed the TRIPOD statement for reporting this clinical prediction model study [19].

Ethics

The study protocol adhered to the ethical guidelines of the 1975 Declaration of Helsinki (6th revision, 2008), as was approved by the ethics committee and the local institutional review board of each participating center. Informed consent was obtained from each patient included in the study.

Results

Patient characteristics

The flowchart of patient inclusion in the derivation and validation cohorts is shown in **Fig. 1 s**. Briefly, of the 1345 eligible patients in the derivation cohort, 1198 (89 %; median [IQR] age, 78 [67–84] years; 626 [52 %] men) were included in the main analysis. Almost half of the patients (n = 588) had a CCI score of at least two; 327 (27 %) were on antiplatelet therapy and 258 (22 %) were on anticoagulation (▶ **Table 1**). The most common presentation was bright red blood per rectum (n = 820; 68 %); the median (IQR) hemoglobin level was 10.3 (8.3–12.3) g/dL and 110 of the patients (9 %) were hemodynamically instable. The bleeding episode occurred in 138 patients who

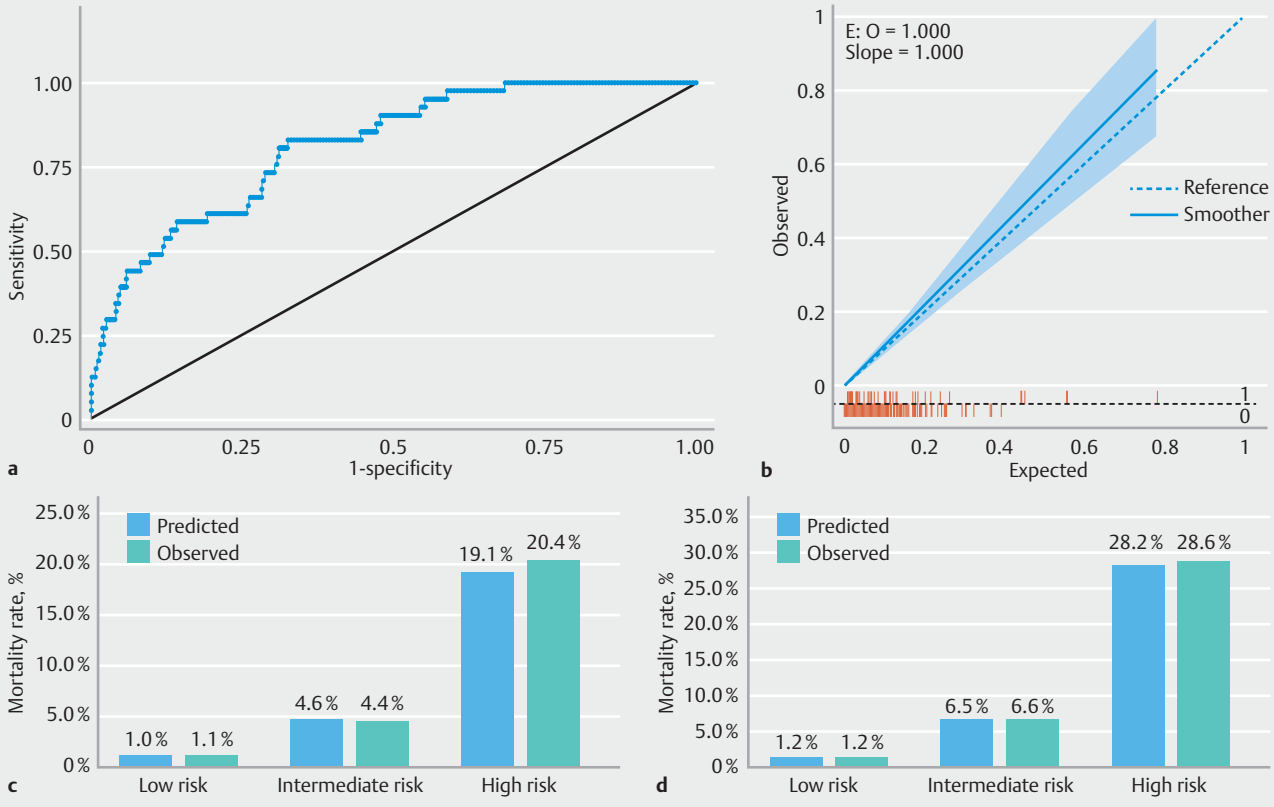
▶ **Table 1** Baseline characteristics, medical history, presentations, and outcomes of the patients included in the derivation and validation cohorts.

	Derivation cohort (n = 1198)	Validation cohort (n = 752)
Age, median (IQR), years	78 (67–84)	76 (66–84)
Sex, male, n (%)	626 (52)	433 (58)
Medical history		
Charlson co-morbidity index, median (IQR)	1 (1–3)	2 (1–4)
▪ Score ≥ 2 points, n (%)	588 (49)	438 (58)
Antiplatelet therapy, n (%)	327 (27)	234 (31)
▪ Dual antiplatelet therapy, n (%)	42 (4)	23 (3)
Anticoagulation therapy, n (%)	258 (22)	204 (27)
▪ Vitamin K antagonists, n (%)	115 (10)	57 (8)
▪ Direct-acting oral anticoagulants, n (%)	143 (12)	147 (20)
Previous admission for LGIB bleeding, n (%)	145 (12)	148 (20)
Clinical presentation		
Type of bleeding		
▪ Bright red blood per rectum, n (%)	820 (68)	442 (59)
▪ Dark blood/clots per rectum, n (%)	195 (16)	235 (31)
▪ Melena, n (%)	183 (15)	75 (10)
Blood on rectal examination, n (%)	N/A	515 (68)
Hemodynamic instability, n (%)	110 (9)	138 (18)
Onset in hospitalized patient, n (%)	138 (12)	197 (26)
Weekend diagnosis, n (%)	126 (11)	190 (25)
Hemoglobin level, median (IQR), g/dL	10.3 (8.3–12.3)	10.1 (8.3–12.6)
Platelet count, median (IQR), $\times 10^9$/L	223 (167–285)	218 (170–288)
INR, median (IQR)	1.1 (1–1.30)	1.1 (1–1.28)
Serum creatine, median (IQR), g/dL	1.1 (0.8–1.4)	0.95 (0.76–1.30)
Outcome		
Need for transfusion, n (%)	515 (43)	329 (44)
Endoscopic treatment, n (%)	208 (17)	159 (21)
Interventional radiology treatment, n (%)	17 (1)	22 (3)
Surgical treatment, n (%)[1]	84 (7)	23 (3)
Rebleeding, n (%)	105 (9)	68 (9)
Death, n (%)	41 (3)	38 (5)

INR, international normalized ratio; LGIB, lower gastrointestinal bleeding; N/A, not available.
[1] Surgery was performed urgently in one case only; in other cases, surgery was elective to treat the cause of the bleeding.

were already hospitalized for other reasons (12 %) and during the weekend in 126 of the patients (11 %).

Colonoscopy was performed in 772 patients (64 %); early (≤ 24 hours) in 197 cases (26 %) and delayed (> 24 hours) in 575 cases (74 %). Only 22 of the patients with hemodynamic instability (20 %) underwent computed tomography angiography (CTA) as their first examination (as per guideline recommendation) [4]. The source of bleeding was identified in 1048 of the patients (87 %), with diverticular disease (n = 257; 21 %), malignancy (n = 147; 12 %), angiodysplasia (n = 124; 10 %), ischemic colitis (n = 115; 10 %), and hemorrhoids (n = 10; 9 %) being the most frequent causes. At the end of follow-up, 105 patients had presented with rebleeding (8.8 %, 95 %CI 7.2 %–10.5 %) and 41 patients had died (3.4 %, 95 %CI 2.5 %–4.6 %).



▶ **Fig. 2** Assessment of the newly developed multivariable model in terms of: **a** discriminative ability (area under the receiver operating curve [AUROC] 0.81 [95%CI 0.75–0.87]); **b** calibration; **c,d** predicted and observed risk of death according to the ALIBI risk groups in: **c** the derivation cohort; **d** the validation cohort.

In the validation cohort, after the exclusion of one patient aged < 18 years and five patients for incomplete data, 752 patients were included in the analysis. The patient demographics and medical histories were comparable between the two cohorts, but patients in the validation cohort were more often hemodynamically instable (18%) and hospitalized at the onset of the LGIB (26%) (▶ **Table 1**).

Colonoscopy was performed in 548 patients (73%), and it was delayed in most cases (n = 402; 73%). Also in this cohort, only 27 of the patients with hemodynamic instability (20%) underwent CTA as their first examination, with comparable outcomes in terms of mortality between the adherent (CTA first) and nonadherent groups ($P = 0.56$). The overall rebleeding and mortality rates in the validation cohort were 9.0% (95%CI 7.1%–11.3%) and 5.1% (95%CI 3.6%–6.9%), respectively.

Further details of patients' diagnoses and treatment, and the etiologies of the LGIB in the two cohorts are reported in **Table 3s**.

Development of a new model to predict mortality

In the univariable analysis, age, CCI, use of NSAIDs, LGIB onset in hospitalized patients, weekend diagnosis, bright red blood at presentation, hemodynamic instability, hemoglobin, INR, and creatine were associated with mortality (**Table 4s**). The final multivariable analysis included age (coefficient 0.08, 95%CI 0.04–0.13; $P < 0.001$), CCI (coefficient 0.16, 95%CI 0.05–0.27; $P = 0.02$), onset in hospitalized patients (coefficient 1.02, 95% CI 0.19–1.86; $P = 0.009$), hemodynamic instability (coefficient 1.19, 95%CI 0.43–1.94; $P = 0.002$), and creatinine (coefficient 0.26, 95%CI 0.01–0.52; $P = 0.04$) (**Table 5s**). The discriminative ability (AUROC 0.81, 95%CI 0.75–0.87) and calibration of the model were excellent (▶ **Fig. 2a,b**).

We then developed a pragmatic scoring system based on these five component variables (▶ **Table 2**): the **A**cute **L**ower gastrointestinal **B**leeding and **I**n-hospital mortality (ALIBI) score. The cumulative score reaches a maximum of 13 points, and the median (IQR) value was 5 (2–6) points. This risk score showed adequate accuracy and calibration (**Fig. 2s**) and the predicted risk for each point is reported in **Table 6s**. Patients could be classified into low risk (0–4 points), intermediate risk (5–9 points), and high risk (≥ 10 points), with a predicted risk of death of 1.0% (95%CI 0.3%–1.6%), 4.6% (95%CI 3.2%–6.0%), and 19.1% (9.4%–28.8%), respectively. The expected and observed mortality rates within these three risk groups are shown in ▶ **Fig. 2c**, with good calibration shown.

▶ **Table 2** The ALIBI score for the prediction of mortality among patients with lower gastrointestinal bleeding (LGIB): **a** score calculation; **b** risk categorization.

a

Variable	Points
Age, years	
▪ ≤70	0
▪ >70 and ≤80	3
▪ >80	5
Charlson co-morbidity index	
▪ 0–1	0
▪ 2–3	1
▪ >3	2
Hemodynamic instability	
▪ No	0
▪ Yes	3
LGIB onset setting	
▪ Out of hospital	0
▪ In hospital	2
Serum creatinine, mg/dL	
▪ <1.5	0
▪ ≥1.5	1
Total score	0–13

b

Risk category	Points	Risk of death
Low	0–4	1%
Intermediate	5–9	5%
High	≥10	19%

Performance of the novel model in the external validation cohort

Using the regression coefficient established in the derivation cohort, we evaluated the diagnostic performance of our multivariable logistic model in an independent cohort of 752 patients. The model showed adequate discrimination (AUROC 0.81, 95%CI 0.74–0.88) and calibration (**Fig. 3s**). The same model was valid also when the numerical score (0–13) was used; discrimination was maintained (AUROC 0.79, 95%CI 0.72–0.86), and the score could estimate the risk of death between 0–50% with excellent calibration (▶ **Fig. 3a,b**). The probability of death in the three risk categories was 1.2% (95% CI 0.3%–2.1%), 6.5% (95%CI 4.7%–8.6%), and 28.4% (95%CI 14.3%–42.4%), with good calibration (▶ **Fig. 2d**).

In the validation cohort, we also evaluated the performance of the Oakland score in predicting mortality in 736 patients (16 patients [2%] had missing data for this score). The model showed adequate accuracy (AUROC 0.73, 95%CI 0.66–0.80), which was numerically lower than the new model (DeLong test, $P=0.15$); however, the model was not well calibrated, as it underestimated the risk of mortality after a certain threshold (▶ **Fig. 3c,d**).

Impact of colonoscopy timing according to the risk strata

We described the mortality rates in patients at low risk (0–4 points) versus high risk (>4 points) and the timing of colonoscopy (▶ **Table 3**). The details of the subpopulation and the methods used for this analysis are reported in **Appendix 2s**. Colonoscopy was defined as early when performed within 24 hours and delayed if performed >24 hours from admission. Patient characteristics in the three groups (no colonoscopy, early colonoscopy, delayed colonoscopy) are reported in **Tables 7s** and **8s**.

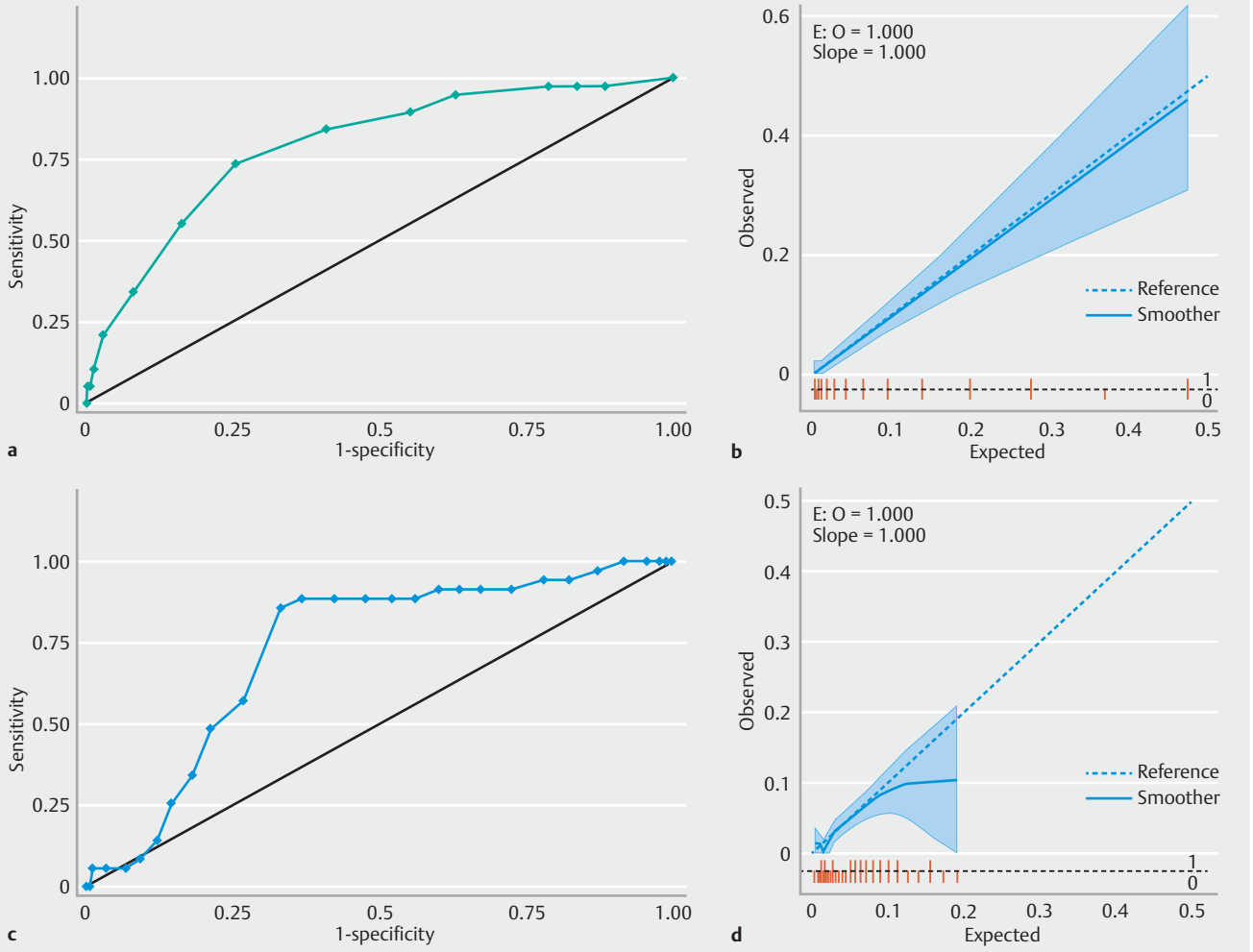
In patients at low risk, mortality rates were 4.8% in patients not undergoing colonoscopy and 0.4% in patients undergoing colonoscopy (adjusted odds ratio [OR] 0.08, 95%CI 0.01–0.88), independently from the timing of the examination. The rate of rebleeding was higher in patients undergoing early colonoscopy (18% vs. 5%; adjusted OR 5.06, 95%CI 1.33–19.25, vs. no colonoscopy).

In high risk patients, mortality rates were 14.0% in patients not undergoing colonoscopy and 6.4% in patients undergoing colonoscopy (adjusted OR 0.36, 95%CI 0.16–0.84); the lower mortality rate was seen only in the delayed colonoscopy group (mortality rate 4.0%; adjusted OR 0.26, 95%CI 0.10–0.71, vs. no colonoscopy). In the early colonoscopy group, patients more frequently underwent an endoscopic hemostatic treatment (47% vs. 22%), but the mortality rate was comparable to that seen in patients who did not undergo colonoscopy at all (12% vs. 13%; adjusted OR: 1.55, 95%CI 0.57–4.21).

Discussion

In this prospective study, we have proposed a novel clinical score that can predict the risk of mortality in patients with acute LGIB. We developed a model based on five readily available variables (age, CCI, setting of bleeding onset, hemodynamic instability, and serum creatine) which was well calibrated and showed good-to-excellent discriminative ability in both the derivation (AUROC 0.81, 95%CI 0.75–0.87) and validation cohorts (AUROC 0.81, 95%CI 0.74–0.88) cohort. We then transformed the model into a simple clinical risk score (the ALIBI score) that stratified the risk of death as low (1%; 0–4 points), intermediate (5%; 5–9 points), or high (19%; 10–13 points) and maintained good discrimination and calibration in both cohorts.

To our knowledge, this is the first prospective study to develop and validate a score that specifically predicts in-hospital mortality in patients with LGIB, so there are no direct comparators. We chose however to assess the prognostic performance of the Oakland score, because it is the most commonly used score in this setting and the only one recommended by current guidelines [4, 20]. The Oakland score was developed to identify



▶ **Fig. 3** Comparison within the external validation cohort for the 736 patients with sufficient data available of: **a,b** the ALIBI score; **c,d** the Oakland score, in terms of: **a,c** discrimination; **b,d** calibration. The area under the curve (AUROC) for the ALIBI score was 0.80 (95 %CI 0.72–0.88) and for the Oakland score was 0.73 (95 %CI 0.66–0.80).

patients at low risk of complications who can be safely discharged [4], but it has also been shown to predict mortality with variable accuracy (AUROC 0.67–0.89) [5, 21, 22]. Our study showed that the Oakland score has acceptable accuracy (AUROC 0.73, 95 %CI 0.66–0.80), but poor calibration for the prediction of in-hospital mortality, making it unsuitable for assessing this outcome. To some extent, this is to be expected as the score was developed to predict a different end point (a composite outcome of absence of rebleeding, blood transfusion, therapeutic intervention, readmission, or death).

Among the recently proposed scores [22, 23, 24], the ABC score was developed to predict mortality in patients with upper gastrointestinal bleeding [23], but also showed promising accuracy in patients with LGIB (AUROC 0.84); as yet, its calibration has not been evaluated and its performance has not been validated in other LGIB cohorts. Unfortunately, we could not assess this score in our patients, because the data on albumin levels and ASA physical status required for the calculation were not routinely collected in our external validation cohort. Future studies are therefore required to compare the prognostic accuracy between the two models.

Several scores have also been previously developed to identify high risk patients, namely to predict "severe bleeding," such as the BLEED [25], NOBLADS [26], and Strate scores [27]; however, these scores used a composite end point as a surrogate for survival and are underpowered or have limited accuracy in predicting mortality in patients with LGIB [11]. Other limitations of such scores include their development only in patients undergoing endoscopy (spectrum bias), and the lack of prospective validation or inconsistent prognostication results, likely owing to the heterogeneity in patient management and outcomes in LGIB. As a result, none of these scores has been introduced into clinical practice or recommended by guidelines [11].

We propose an easy, reliable, and informative tool that can estimate the risk of in-hospital mortality during admission for an episode of LGIB. Our score was able to accurately predict the likelihood of an uncommon, but not negligible, event such

▶ **Table 3** Mortality rate according to the new risk score and the timing of endoscopy (early [≤ 24 hours] vs. delayed [> 24 hours]) in the validation cohort.

Risk category (score)	Colonoscopy performed	Mortality rate	Rebleeding rate	Endoscopic hemostatic treatment	Embolization/ surgery
Low (0–4)	No (n = 63)	4.8 %	4.8 %	0 %	3.2 %
	Yes (n = 249)[1]	0.4 %	8.4 %	24.9 %	5.6 %)
	▪ Early (n = 71)	0 %	18.3 %	47.9 %	7.0 %
	▪ Delayed (n = 174)	0.6 %	4.6 %	16.1 %	4.6 %
Intermediate or high (≥ 5)	No (n = 103)	12.6 %	7.8 %	0 %	0.1 %
	Yes (n = 312)[2]	6.4 %	11.2 %	27.6 %	5.1 %
	▪ Early (n = 75)	12.0 %	14.7 %	46.7 %	10.7 %
	▪ Delayed (n = 227)	4.0 %	9.7 %	21.6 %	3.5 %

[1] Timing of endoscopy missing for four patients.
[2] Timing of endoscopy was missing for 10 patients.

as death in LGIB, which on the other hand remains a frequent, burdensome, and costly condition.

The population at risk includes patients admitted with LGIB who are generally outside the safe discharge criteria according to the Oakland score, yet represent around 90 % of patients with this condition [5, 9]. These patients are often frail, and the burden of co-morbidities is relevant [28], so care is generally nuanced, patient-centered, and individualized. In this complex setting, a score that reflects both the severity of the bleeding episode and the general condition of the patient may better evaluate prognosis and inform clinicians, patients, and their families in the decision-making process.

Regarding the timing of endoscopy, data from randomized trials have failed to show the benefit of early colonoscopy in terms of improved clinical outcomes, and current guidelines recommend performing a colonoscopy sometime during admission [4, 29,30]. A universal strategy may not however fit all patients with LGIB, and the extent to which the severity of the bleeding episode influences the harms and benefits of an early colonoscopy is largely unknown. In our study, we benchmarked the mortality rates according to the predicted risk (low vs. intermediate/high), the choice/possibility to undergo a colonoscopy, and its timing (early vs. delayed). In patients at intermediate or high risk, the delayed colonoscopy group was associated with a lower mortality rate, even after adjustment for the severity of the bleeding episode. These observations could reflect the importance of patient stabilization and supportive care in this category of patients [31], but may also be subject to residual confounding, as no cause–effect relationship can be assumed from our study design. Therefore, future studies and trials are needed to investigate whether a score identifying high risk patients can contribute to personalized management in LGIB.

At an institutional level, the score can be useful for auditing purposes or monitoring local outcomes over different periods of time.

Finally, only 20 % of patients with hemodynamic instability underwent a CTA as their first examination, suggesting suboptimal adherence to guidelines. The implications of nonadherence in terms of prognosis go beyond the scope of the present study however and deserve further investigation.

In the research field, our study provides a robust framework for the assessment of prognosis in patients with LGIB. It can be used to assess the interactions between disease severity, diagnostic work-up, treatment, and clinical outcomes in observational studies or to design future clinical trials in this field. For instance, focusing on high risk patients could reduce the sample size required to evaluate the potential benefit of diagnostic or therapeutic interventions. Future prospective studies conducted by independent investigators should provide further external validation of the prognostic model.

The main strengths of our study lie in its prospective design, the sample size, and the external validation. The multicenter and international nature of the validation cohort supports the generalizability of our findings. A limitation is the lack of longer follow-up data on mortality; however, post-discharge events could be influenced more by age and co-morbidities rather than the bleeding episodes and introduce a bias. Moreover, we did not compare the performance of our model to that of other published scores, with the exception of the Oakland score. Finally, the observational nature of the study does not allow establishment of a causative relationship between the timing of colonoscopy and mortality in high risk patients, so these data should be interpreted with caution and future randomized controlled trials are required to address this issue.

In conclusion, we propose a new pragmatic risk score (the ALIBI score), based on variables that are readily available at diagnosis, that can predict the risk of in-hospital mortality in patients with acute LGIB. The risk categories can be used to inform patients and support medical decisions regarding diagnostic work-up and treatment of patients with this condition.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Peery AF, Crockett SD, Murphy CC et al. Burden and cost of gastrointestinal, liver, and pancreatic diseases in the united states: update 2021. Gastroenterology 2022; 162: 621–644doi:10.1053/j.gastro.2021.10.017

[2] Lanas A, García-Rodríguez LA, Polo-Tomás M et al. The changing face of hospitalisation due to gastrointestinal bleeding and perforation. Aliment Pharmacol Ther 2011; 33: 585–591doi:10.1111/J.1365-2036.2010.04563.X

[3] Williams JG, Roberts SE, Ali MF et al. Gastroenterology services in the UK. The burden of disease, and the organisation and delivery of services for gastrointestinal and liver disorders: a review of the evidence. . Gut 2007; 56: (Suppl. 01): 1–113doi:10.1136/gut.2006.117598

[4] Triantafyllou K, Gkolfakis P, Gralnek IM et al. Diagnosis and management of acute lower gastrointestinal bleeding: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy 2021; 53: 850–868doi:10.1055/A-1496-8969

[5] Oakland K, Jairath V, Uberoi R et al. Derivation and validation of a novel risk score for safe discharge after acute lower gastrointestinal bleeding: a modelling study. Lancet Gastroenterol Hepatol 2017; 2: 635–643doi:10.1016/S2468-1253(17)30150-4

[6] Strate LL, Ayanian JZ, Kotler G et al. Risk factors for mortality in lower intestinal bleeding. Clin Gastroenterol Hepatol 2008; 6: 1004–1010doi:10.1016/j.cgh.2008.03.021

[7] Lanas A, García-Rodríguez LA, Polo-Tomás M et al. Time trends and impact of upper and lower gastrointestinal bleeding and perforation in clinical practice. Am J Gastroenterol 2009; 104: 1633–1641doi:10.1038/ajg.2009.164

[8] Radaelli F, Rocchetto S, Piagnani A et al. Scoring systems for risk stratification in upper and lower gastrointestinal bleeding. Best Pract Res Clin Gastroenterol 2023; 67: 101871doi:10.1016/j.bpg.2023.101871

[9] Oakland K, Kothiwale S, Forehand T et al. External validation of the Oakland score to assess safe hospital discharge among adult patients with acute lower gastrointestinal bleeding in the US. JAMA Netw Open 2020; 3: E209630doi:10.1001/jamanetworkopen.2020.9630

[10] Tapaskar N, Jones B, Mei S et al. Comparison of clinical prediction tools and identification of risk factors for adverse outcomes in acute lower GI bleeding. Gastrointest Endosc 2019; 89: 1005–1013.e2doi:10.1016/j.gie.2018.12.011

[11] Almaghrabi M, Gandhi M, Guizzetti L et al. Comparison of risk scores for lower gastrointestinal bleeding: a systematic review and meta-analysis. JAMA Netw Open 2022; 5: E2214253doi:10.1001/jamanetworkopen.2022.14253

[12] Radaelli F, Frazzoni L, Repici A et al. Clinical management and patient outcomes of acute lower gastrointestinal bleeding. A multicenter, prospective, cohort study. . Dig Liver Dis 2021; 53: 1141–1147doi:10.1016/j.dld.2021.01.002

[13] Box GEP, Tidwell PW. Transformation of the independent variables. Technometrics 1962; 4: 531–550doi:10.1080/00401706.1962.10490038

[14] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; 44: 837–845

[15] Steyerberg EW, Harrell FE, Borsboom GJJM et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001; 54: 774–781doi:10.1016/S0895-4356(01)00341-9

[16] Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman and Hall/CRC; 1994: doi:10.1201/9780429246593

[17] Iba K, Shinozaki T, Maruo K et al. Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. BMC Med Res Methodol 2021; 21: 9doi:10.1186/S12874-020-01201-W

[18] Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004; 23: 1631–1660doi:10.1002/sim.1742

[19] Moons KGM, Altman DG, Reitsma JB et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 2015; 162: W1–W73doi:10.7326/M14-0698

[20] Sengupta N, Feuerstein JD, Jairath V et al. Management of patients with acute lower gastrointestinal bleeding: an updated ACG Guideline. Am J Gastroenterol 2023; 118: 208–231doi:10.14309/ajg.0000000000002130

[21] Prasitvarakul K, Attanath N, Chang A. Comparison of scoring systems for predicting clinical outcomes of acute lower gastrointestinal bleeding: A prospective cohort study. World J Surg 2024; 48: 474–483doi:10.1002/wjs.12053

[22] Shung DL, Chan CE, You K et al. Validation of an electronic health record-based machine learning model compared with clinical risk scores for gastrointestinal bleeding. Gastroenterology 2024; 167: 1198–1212doi:10.1053/j.gastro.2024.06.030

[23] Laursen SB, Oakland K, Laine L et al. ABC score: a new risk score that accurately predicts mortality in acute upper and lower gastrointestinal bleeding: an international multicentre study. Gut 2021; 70: 707–716doi:10.1136/gutjnl-2019-320002

[24] Sengupta N, Tapper EB. Derivation and internal validation of a clinical prediction tool for 30-day mortality in lower gastrointestinal bleeding. Am J Med 2017; 130: 601.e1–e601.e8doi:10.1016/j.amjmed.2016.12.009

[25] Kollef MH, O'Brien JD, Zuckerman GR et al. BLEED: a classification tool to predict outcomes in patients with acute upper and lower gastrointestinal hemorrhage. Crit Care Med 1997; 25: 1125–1132doi:10.1097/00003246-199707000-00011

[26] Aoki T, Nagata N, Shimbo T et al. Development and validation of a risk scoring system for severe acute lower gastrointestinal bleeding. Clin Gastroenterol Hepatol 2016; 14: 1562–1570.e2doi:10.1016/j.cgh.2016.05.042

[27] Strate LL, Orav EJ, Syngal S. Early predictors of severity in acute lower intestinal tract bleeding. Arch Intern Med 2003; 163: 838–843doi:10.1001/archinte.163.7.838

[28] Oakland K, Guy R, Uberoi R et al. Acute lower GI bleeding in the UK: patient characteristics, interventions and outcomes in the first nationwide audit. Gut 2018; 67: 654–662doi:10.1136/gutjnl-2016-313428

[29] Niikura R, Nagata N, Yamada A et al. Efficacy and safety of early vs elective colonoscopy for acute lower gastrointestinal bleeding. Gastroenterology 2020; 158: 168–175.e6doi:10.1053/j.gastro.2019.09.010

[30] Laine L, Shah A. Randomized trial of urgent vs. elective colonoscopy in patients hospitalized with lower GI bleeding. . Am J Gastroenterol 2010; 105: 2636–2641doi:10.1038/ajg.2010.277

[31] Lau JYW, Yu Y, Tang RSY et al. Timing of endoscopy for acute upper gastrointestinal bleeding. NEJM 2020; 382: 1299–1308doi:10.1056/nejmoa1912484