ORIGINAL ARTICLE DOI: https://doi.org/10.18502/fem.v8i3.16331

Comparative analysis of four upper gastrointestinal bleeding scoring systems for predicting multiple outcomes: an observational study in the emergency department

Najmeh Tohidi[1], Mitra Movahedi[1], Mohammad Rezaei Zadeh Rukerd[2], Hanieh Mirkamali[3], Seyed Danial Alizadeh[4], Mohammad Javad Najafzadeh[5], Amin Honarmand[1], Mehran Ilaghi[6], Pouria Pourzand[7], Amirhossein Mirafzal[1]***

1. Department of Emergency Medicine, Kerman University of Medical Sciences, Kerman, Iran.
2. Gastroenterology and Hepatology Research Center, Institute of Basic and Clinical Physiology Sciences, Kerman University of Medical Sciences, Kerman, Iran.
3. Student Research Committee, School of Medicine, Kerman University of Medical Sciences, Kerman, Iran.
4. Faculty of Medicine, Kerman University of Medical Sciences, Kerman, Iran.
5. Department of Orthopedic Surgery, Kerman University of Medical Sciences, Kerman, Iran.
6. Institute of Neuropharmacology, Kerman Neuroscience Research Center, Kerman University of Medical Sciences, Kerman, Iran.
7. Department of Emergency Medicine, School of Medicine, University of Minnesota, Minneapolis, United States.

***Primary corresponding author:** Amirhossein Mirafzal; **Email:** a.mirafzal@yahoo.com
****Secondary corresponding author:** Pouria Pourzand; **Email:** pourzand@umn.edu

Published online: 2024-07-17

Abstract: **Objective:** Numerous scoring systems have been developed to assess the risk associated with upper gastrointestinal bleeding (UGIB), and several studies have investigated their comparative accuracy in predicting patient outcomes. This study was undertaken to compare four well-known scoring systems, namely the pre-endoscopy Rockall score, full Rockall score, Glasgow-Blatchford Bleeding score (GBS), and AIMS65, with the aim of predicting five distinct outcomes in cases of non-variceal UGIB.

Methods: This prospective observational study was conducted focusing on adult patients with UGIB presenting to the emergency department (ED). The primary endpoints of this study included in-hospital mortality, the need for re-endoscopy, the requirements for packed red blood cell (PRBC) transfusion, massive transfusion, and one-month rebleeding.

Results: A total number of 320 patients were enrolled, with 44 (13.75%) in-hospital deaths. Based on the area under the curves (AUC), while certain scores outperformed others in specific outcome prediction, the AIMS65 scoring system demonstrated superior predictive capability for both in-hospital mortality (0.91) and massive transfusion (0.71). Regarding PRBC transfusion requirements, both AIMS65 and GBS exhibited similar predictive capacities (AUC=0.67 and 0.68, respectively). In terms of re-endoscopy and one-month rebleeding, the GBS scoring system displayed slightly better performance compared to the other systems (AUC=0.61 and 0.63, respectively). In the composite outcome, all scores had significant associations, and among them, the AIMS-65 score had the highest AUC (0.76).

Conclusion: The AIMS65 scoring system was the most reliable tool for predicting in-hospital mortality and, to a lesser extent, massive transfusion requirements, while GBS and AIMS65 could be moderately and cautiously relied on for preparations regarding the need for PRBC transfusion.

Keywords: Outcomes; Prediction; Scoring Systems; Upper Gastrointestinal Bleeding

Cite this article as: Tohidi N, Movahedi M, Rezaei Zadeh Rukerd M, et al. Comparative analysis of four upper gastrointestinal bleeding scoring systems for predicting multiple outcomes: an observational study in the emergency department. Front Emerg Med. 2024;8(3):e24.

1. Introduction

Upper gastrointestinal bleeding (UGIB), characterized by blood loss from the esophagus, stomach, or duodenum in close proximity to the Treitz ligament, is a life-threatening condition with high morbidity within the emergency department (ED) setting (1-3). Although UGIB has been associated with an average mortality rate of 10%, the evidence on recent progress in reducing mortality rates of UGIB remains inconclusive (4-7). Proper management of UGIB could include risk assessment of outcomes as an initial step to address the high complication and mortality rates (8).

Risk assessment of UGIB is commonly conducted utilizing diverse scoring systems and models, which have been

introduced, evaluated, and compared within the existing medical literature (9,10). These scoring systems include those employed prior to the endoscopic study, such as pre-endoscopy Rockall, AIMS65, and Glasgow-Blatchford bleeding score (GBS), as well as those reliant on endoscopy, including full Rockall, Baylor bleeding score, The Progetto Nazionale Emorragia Digestiva, and Cedars-Sinai Medical Center predictive index (8-10). For the purpose of comparison, most studies are focused on a few of those systems, namely Rockall, GBS, and AIMS65 (9,11). Nevertheless, within the current medical literature, there is a notable absence of a scoring system that can precisely predict all critical clinical outcomes associated with UGIB (12). International consensus groups have recommended using GBS, using a threshold of ≤1, to identify patients at a very low risk of complications (13,14). However, despite their recommendations supporting the use of GBS, the literature presents varying and conflicting results when it comes to predicting adverse outcomes such as mortality (14).

In the current study, we aimed to determine the most appropriate scoring system among the four commonly employed measures, namely pre-endoscopy and full Rockall score, GBS, and AIMS65, for predicting five critical outcomes: in-hospital mortality, the need for re-endoscopy, (packed red blood cell) PRBC transfusion requirements, massive transfusion, and the rate of one-month rebleeding.

2. Methods

2.1. Study design

This observational study was performed in the ED of an academic referral hospital with an annual ED census of approximately 40,000 in a 6-month period from April to October 2022. At our hospital's ED, patients are triaged based on the five-level emergency severity index triage system. For patients triaged to emergency severity index levels 1 to 4, first-line visits are performed by a resident and an attending physician of emergency medicine (EM), followed by expert consultations as required at the discretion of the EM service. Patients with the complaint of UGIB are triaged to levels 1 to 3 (mostly 1 or 2) depending on the predicted required facilities, risk status, and vital signs.

This research protocol underwent a rigorous review process and received formal approval from both the Ethics Committee and the Institutional Review Board of Kerman University of Medical Sciences, bearing the ethical code (IR.KMU.AH.REC.1399.002). Every aspect of our methodology strictly adheres to pertinent guidelines and regulatory frameworks. Prior to their participation, all individuals provided written informed consent following a comprehensive explanation of the study's objectives and procedures.

2.2. Study population, variables, and protocol

The primary aim of this study is to predict in-hospital mortality by the scoring systems such as pre-endoscopy and full Rockall score, GBS, and AIMS65. The secondary outcomes are predicting the need for re-endoscopy, PRBC transfusion requirements, massive transfusion, and the rate of one-month rebleeding with the mentioned scoring systems.

The study included all adult patients (age >18 years) who presented to the ED with suspected UGIB and exhibited symptoms such as hematochezia, rectorrhagia, or melena within the designated study period and subsequently underwent endoscopy as convenience sampling.

Exclusion criteria were patients with diagnoses other than UGIB, those subsequently diagnosed with variceal hemorrhage, those who did not undergo endoscopy for any reason, and patients who declined to participate or were lost to follow-up. Additionally, patients who were deemed unlikely to have UGIB or its associated mortality risks, such as those with sepsis or respiratory failure, were also excluded based on consensus agreement.

The main study variables included demographic characteristics, background diseases, initial vital signs, hemodynamic shock status, endoscopic diagnosis, hemoglobin levels, blood urea nitrogen (BUN) levels, albumin levels, and altered mental status. These variables were essential for the computation of the pre-endoscopy Rockall score, full Rockall score, GBS, and AIMS65 score. The calculation of the aforementioned scores was carried out by two senior residents of EM. To evaluate the agreement between the two evaluators, the kappa coefficient was utilized as a measure of inter-observer agreement. Where there was disagreement, a third evaluator, an attending physician of EM, resolved the discrepancies. The scoring systems were explained below:

1) Glasgow-Blatchford bleeding score: The GBS was calculated based on predetermined gender, clinical, and laboratory parameters, including hemoglobin levels, BUN, initial systolic blood pressure, pulse rate, the presence of melena or recent syncope, and the presence of hepatic or cardiac disease. The GBS ranges from 0 to 23, and a higher GBS is also correlated with a higher likelihood of needing intervention (scores ≥6 are associated with >50% risk of needing intervention) (15), 2) AIMS65 score: the score was calculated based on five clinical parameters: albumin level, international normalized ratio (INR), altered mental status, systolic blood pressure, and age. Each parameter is assigned a score of 0 or 1, with a total score ranging from 0 to 5. Higher scores indicate increased in-hospital mortality risk (16), 3) pre-endoscopy Rockall score: this score was calculated for each patient using predetermined criteria, including age, presence of shock, comorbidities, and clinical signs of recent bleeding. This score ranges from 0 to 7, with higher scores indicating increased severity and a worse prognosis (17), and 4) full Rockall score: this score incorporates endoscopic findings in addition to the pre-endoscopy criteria used in the pre-endoscopy Rockall score. Endoscopic findings such as dark spots, blood in the upper gastrointestinal tract, adherent clots, and visible or spurting vessels are included in the calculation. The full Rockall score ranges from

0 to 11, with higher scores indicating greater rebleeding and mortality risk (17).

All patients included in the study were closely monitored during their entire in-hospital admission period. Additionally, a follow-up was conducted after one month via telephone interviews to assess five specific outcomes: in-hospital mortality, the need for re-endoscopy during the hospital admission, requirements for PRBC transfusion, the need for massive transfusion (defined as receiving more than 10 units of PRBCs within a 24-hour period), and the rate of rebleeding within one month.

2.3. Sample size

A minimum sample size of 260 (with at least 26 in-hospital deaths) was calculated according to the formula of sample size calculation for diagnostic tests, considering respective type I and II errors of 5 and 20%, a mortality rate of 10%, and the receiver operating characteristics curve (ROC) reported in a similar study (18).

2.4. Statistical analysis

Quantitative variables with a normal distribution were described using the mean and standard deviation, while variables with a non-normal distribution were described using the median and interquartile range. Qualitative variables were presented as percentages or frequencies. The normal distribution of quantitative variables was assessed according to the Kolmogorov-Smirnov test. An independent t-test was employed for the comparison of normally distributed variables among groups. Otherwise, a Mann-Whitney test was utilized. The ROC curve was applied to illustrate the area under curve (AUC) with 95% confidence intervals (CI), sensitivity, specificity, positive predictive value, and negative predictive value of specific cutoff points for predicting the outcomes. The cutoff values of each test were selected separately for each outcome using Youden's index method. The Youden index, a commonly employed technique in diagnostic test evaluation, maximizes the sum of sensitivity and specificity to identify the optimal threshold for distinguishing between positive and negative outcomes (19). The positive likelihood ratio (PLR) is computed by dividing the true positive rate (sensitivity) by the false positive rate (1specificity). In contrast, the negative likelihood ratio (NLR) is obtained by dividing the false negative rate (1sensitivity) by the true negative rate (specificity) (20).

A P-value of less than 0.05 was considered statistically significant in all statistical tests. Statistical package for social sciences version 16.0 (SPSS Inc., Chicago, IL, USA) was used for analysis. ROC curves were plotted, and the corresponding AUC values were calculated using R (version 4.2.1). Comparison among ROC curve AUCs was done by using the Hanley-McNeil method to calculate z scores between each two curves.

3. Results

3.1. Basic characteristics

A total of 320 patients were included in the study, with male-to-female ratio of 1.4:1. The enrollment process flow diagram has been depicted in figure 1. Among the patients, heart failure, chronic kidney disease, liver cirrhosis, and a history of malignancy were observed in 74 (23.12%), 51 (15.93%), 12 (3.75%), and 29 (9.06%) cases, respectively. The descriptive statistics of quantitative variables are presented in table 1.

The clinical presentation of UGIB featured melena (76.25%), hematemesis with fresh blood (32.18%), hematemesis with coffee ground appearance (24.68%), and hematochezia (7.50%) within our study population. Additionally, 118 (36.87%) cases exhibited an altered mental status. Outcomes observed among the patients were as follows: in-hospital mortality (n= 44, 13.75%), re-endoscopy requirement (n= 84, 26.25%), received PRBC transfusion (n=132, 41.25%), needed massive transfusion (n= 22, 6.87%), and experienced one-month rebleeding (n=44, 13.75%). Regarding agreements between the two evaluators, kappa coefficient values were 0.90, 0.90, 0.95 and 0.85 for pre-endoscopy Rockall, Rockall, AIMS-65 and GBS, respectively.

3.2. Receiver operating curve (ROC) analysis

The results of univariate analysis indicated that, with the exception of AIMS65 and pre-endoscopy Rockall for re-endoscopy, all scores exhibited a statistically significant association with the outcomes. Table 2 presents a comprehensive overview of each scoring system for predicting outcomes and composite outcome. It provides detailed information on the performance of each scoring system for the prediction of the specified outcomes. Additionally, figure 2 displays the ROC curve for GBS, AIMS65, pre-endoscopy Rockall score, and full Rockall score as predictors of in-hospital mortality. Comparison of ROC curves showed that AIMS65 score was superior to GBS (z= 3.14), pre-endoscopy Rockall (z= 2.56), and Rockall (z=0.3.73) scoring systems for the prediction of in-hospital mortality.

4. Discussion

In the current study, four established scoring systems, namely the pre-endoscopy Rockall score, full Rockall score, GBS, and AIMS65, were evaluated and compared regarding their performance across five outcomes in non-variceal UGIB patients, including in-hospital mortality, re-endoscopy need, PRBC transfusion requirements, massive transfusion, and one-month rebleeding. The results from this study highlighted the predictive abilities of AIMS65 for mortality, the pre-endoscopy Rockall score for re-endoscopy, and the GBS for re-endoscopy and rebleeding within one month, while both AIMS65 and GBS displayed similar predictive capacities for PRBC transfusion requirements. In the composite outcome, all scores had significant associations, and among them, the AIMS-65 score had the highest AUC, specificity,

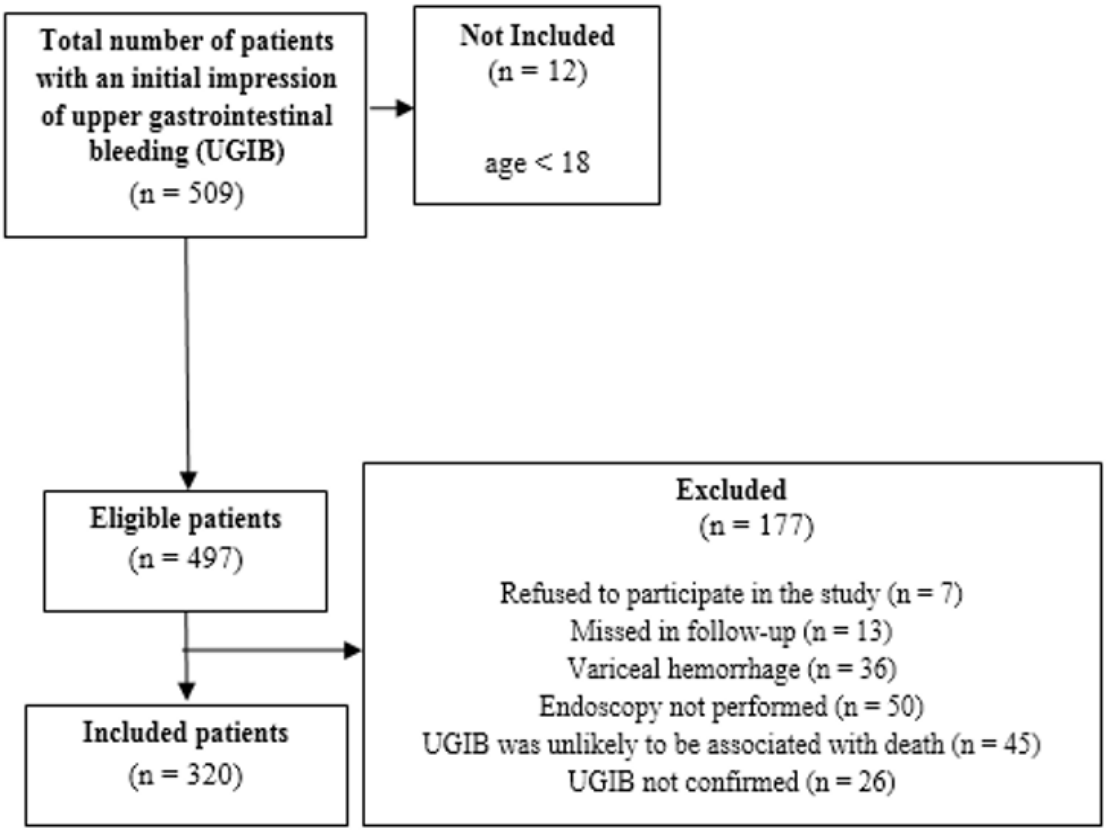

Figure 1 Flow diagram of the enrollment process

Table 1 Characteristics and proportion of patients studied for comparative analysis of four scoring systems

Scoring systems	Glasgow-Blatchford Bleeding score, median (IQR)	13 (5)
	AIMS65 score, median (IQR)	2 (3)
	Pre-endoscopic Rockall score, median (IQR)	3 (2)
	Full Rockall score, median (IQR)	5 (4)
Basic parameters	Age (year), median (IQR)	59·5 (29)
	Sex (male), n (%)	186 (58.1)
	SBP (mmHg), median (IQR)	110 (35)
	Heart rate (beats/min), median (IQR)	92 (27)
	Hemoglobin (mg/dL), median (IQR)	8.5 (2.2)
	BUN (mg/dL), median (IQR)	52 (51)
	Base deficit (mmol/L), median (IQR)	3 (4)
	INR, median (IQR)	1.5 (0·5)
	Serum albumin (g/dL), median (IQR)	3 (0·5)

BUN: Blood urea nitrogen; INR: International normalized ratio; IQR: Interquartile range; N: Number; SBP: Systolic blood pressure

and PLR. Our results firmly support the notion that AIMS65 holds a preeminent role as the predictor of major outcomes combined, surpassing the other evaluated scoring systems by a significant margin, affirming its unrivaled effectiveness in clinical prognostication.

Up until now, multiple comparative studies have been conducted between the Rockall score and the GBS. One of them found that the full Rockall score tended to outperform the GBS in predicting one-month mortality. On the other hand, GBS demonstrated better predictive ability for rebleeding and PRBC transfusion requirements (21). In a separate study, the prognostic accuracy of the GBS was found to surpass that of the pre-endoscopy Rockall score while being comparable to the full Rockall score. However, despite these findings, the AUC values of these scoring systems were deemed insufficient to warrant their exclusive recommendation for routine use in clinical practice (22). Also, recently performed research showed that Rockall was more reliable than GBS for the prediction of mortality and endoscopic requirements, while GBS was more accurate for the prediction of transfu-

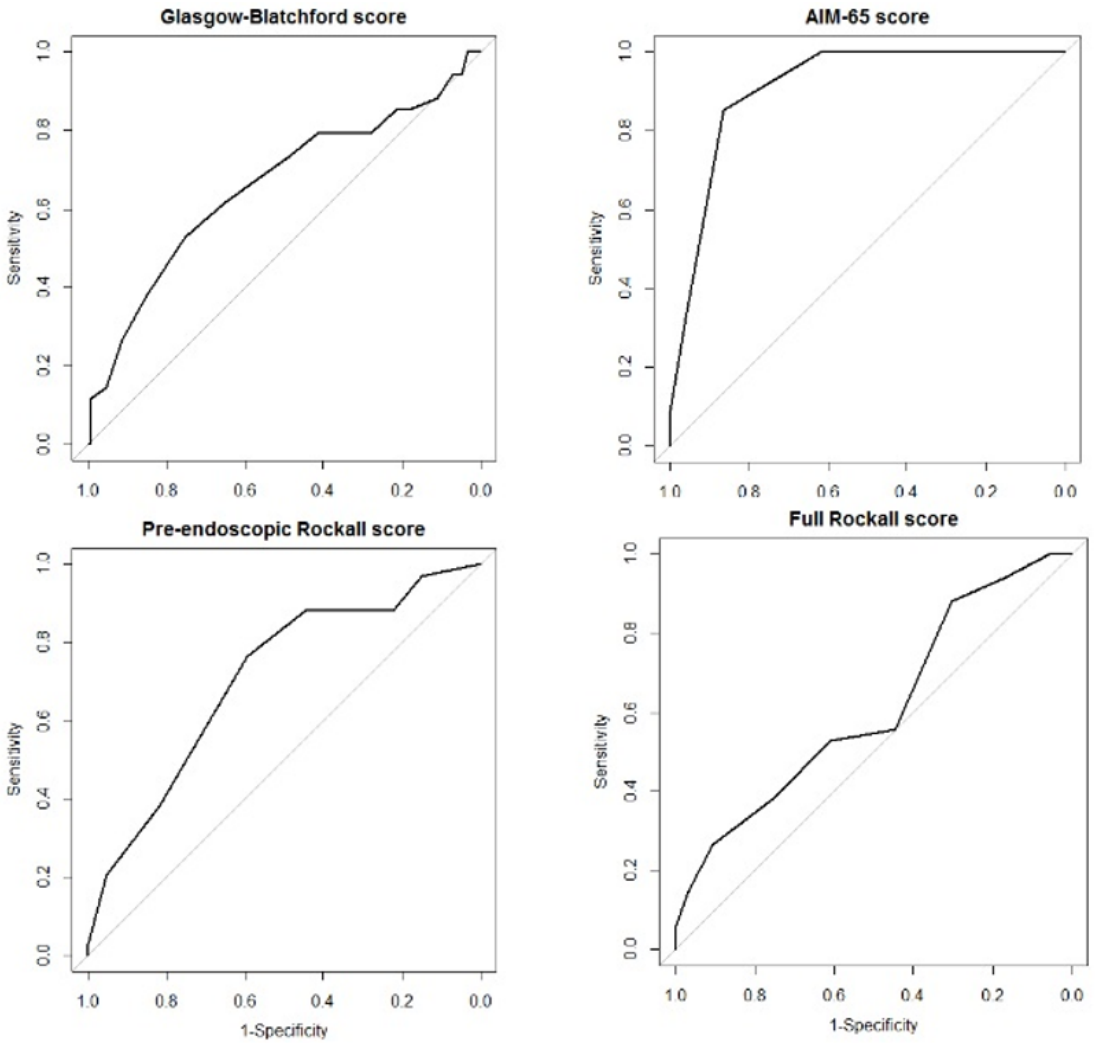

Figure 2 Receiver operating characteristics curves (ROC) for prediction of in-hospital mortality by Glasgow-Blatchford score (AUC=0·66, 95% CI: 0·54,0·75), AIMS65 score (AUC=0·91, 95% CI: 0·86,0·95), pre-endoscopic Rockall score (AUC= 0·71, 95% CI: 0·61,0·80), and full Rockall score (AUC= 0·61, 95% CI: 0·5,0·72)

sion requirements (11). Bryant et al.'s study further supports the superiority of the GBS over the full Rockall score in predicting transfusion requirements. According to their findings, GBS outperformed the full Rockall score specifically in predicting the need for transfusion. However, the performance of the two scoring systems was deemed equal for the other outcomes assessed in their study (23). Due to the limitations seen in current scoring systems' prognostic accuracy, like insufficient AUC values for the pre-endoscopy Rockall score and full Rockall score and the variable performance of the GBS across different outcomes, there is a strong demand for a more reliable and comprehensive predictive test.

Recently, the introduction of AIMS65 into the comparisons between scoring systems has indeed brought about alterations in the results and recommendations. According to a study conducted by Robertson et al. involving 424 patients, AIMS65 demonstrated greater accuracy in predicting mortality, the need for intensive care unit admission, and length of hospital stay compared to GBS and the full Rockall score. However, GBS retained its superiority in predicting transfusion requirements (24). In another study involving 309 patients, the three scoring systems (GBS, full Rockall score, and AIMS65) were found to be comparable in their predictive abilities. GBS showed a slight advantage in predicting the need for transfusion, while AIMS65 demonstrated some advantage in predicting delayed mortality. However, it is worth noting that the reported AUC values for these systems fell within a range indicating moderately acceptable prediction quality (25). Nevertheless, there are research studies that have demonstrated a significant advantage of AIMS65 over other scoring systems, particularly in predicting mortality outcomes. These studies have reported AUC values for AIMS65 in the ranges indicating good predictive capability (18,26-30).

Although all these scoring systems include indicators of advanced hepatic disease, they have been shown to be of

Table 2 Associations between prognostic scores and outcomes (including 95% confidence intervals) with underscored values denoting the best prognostic score's area under the curve for each outcome

		Glasgow-Blatchford score	AIMS65 score	Pre-endoscopic Rockall score	Full Rockall score
In-hospital mortality	Cutoff value*	14	3	4	4
	AUC (95% CI)	0.66 (0.54,0.75)	0.91 (0.86,0.95)	0·71 (0.61,0.80)	0·61 (0.50,0.72)
	Sensitivity† (95% CI)	61 (43,77)	85 (68,95)	76 (58,89)	88 (72,96)
	Specificity† (95% CI)	66 (56,73)	86 (79,91)	59 (50,68)	30 (22,38)
	PPV† (95% CI)	32 (25,40)	63 (5173)	33 (27,40)	25 (22,28)
	NPV† (95% CI)	86 (80,90)	95 (90,97)	90 (83,94)	90 (78,96)
	PLR (95% CI)	1·77 (1.24,2.53)	6·27 (3.95,9.97)	1·89 (1.42,2.51)	1.26 (1.07,1.49)
	NLR (95% CI)	0·59 (0.38,0.92)	0·17 (0.08,0.38)	0·40 (0.21,0.74)	0.39 (0.15,1.02)
	Accuracy† (95% CI)	64 (56,71)	86 (79,91)	63 (55,70)	42 (34,50)
Re-endoscopy	Cutoff value	11	1	2	5
	AUC (95% CI)	0·61 (0.52,0.70)	0.56 (0.47,0.65)	0.54 (0.45,0.64)	0.60 (0.51-0.70)
	Sensitivity† (95% CI)	90 (77,97)	92 (80,98)	92 (80,98)	75 (60,87)
	Specificity† (95% CI)	32 (23,41)	32 (23,41)	24 (17,33)	52 (42,60)
	PPV† (95% CI)	32 (28,35)	32 (28,35)	30 (27,33)	35 (30,41)
	NPV† (95% CI)	90 (78,96)	92 (80,97)	90 (75,96)	85 (77,91)
	PLR (95% CI)	1·33 (1.14,1.56)	1.37 (1.18,1.59)	1.23 (1.08,1.41)	1.58 (1.23,2.03)
	NLR (95% CI)	0.30 (0.11,0.78)	0.23 (0.07,0.70)	0.29 (0.09,0.90)	0.46 (0.26,0.81)
	Accuracy† (95% CI)	47 (39,55)	47 (39,55)	42 (37,50)	58 (50,65)
Packed red blood cell transfusion	Cutoff value	12	2	3	5
	AUC (95% CI)	0.68 (0.59,0.77)	0.67 (0.58,0.76)	0.61 (0.52,0.71)	0.65 (0.56,0.75)
	Sensitivity† (95% CI)	72 (63,80)	60 (51,70)	68 (58,76)	61 (52,71)
	Specificity† (95% CI)	60 (44,73)	70 (54,81)	51 (36,65)	60 (44,73)
	PPV† (95% CI)	80 (73,85)	81 (74,87)	76 (69,81)	77 (70,83)
	NPV† (95% CI)	49 (39,58)	44 (36,51)	41 (32,51)	40 (33,49)
	PLR (95% CI)	1.79 (1.25,2.55)	1.99 (1.27,3.11)	1.40 (1.02,1.91)	151 (1.06-2.20)
	NLR (95% CI)	0.46 (0.31,0.67)	0.56 (0.42,0.76)	0.62 (0.4,-0.91)	0.65 (0.46,0.89)
	Accuracy† (95% CI)	68 (60,75)	63 (55,71)	63 (55,70)	61 (53,68)
Massive transfusion	Cutoff value	12	2	3	6
	AUC (95% CI)	0·61 (0.52,0.71)	0.71 (0.63,0.80)	0.60 (0.49,0.70)	0.60 (050–,0.70)
	Sensitivity† (95% CI)	80 (65,91)	75 (59,87)	73 (57,86)	54 (38,70)
	Specificity† (95% CI)	44 (34,52)	57 (47,65)	43 (32,50)	63 (53,71)
	PPV† (95% CI)	33 (29,38)	37 (31,44)	31 (26,36)	34 (26,42)
	NPV† (95% CI)	86 (76,92)	87 (79,92)	81 (71,88)	79 (73,84)
	PLR (95% CI)	1·43 (1.15,1.77)	1.75 (1.34,2.29)	1.26 (1.00,1.60)	147 (1.02,2.11)
	NLR (95% CI)	0.44 (0.230.85)	0·.43 (0.25,0.75)	0.63 (0.36,1.09)	0.72 (0.50,1.03)
	Accuracy† (95% CI)	53 (45,61)	61 (53,69)	50 (42,58)	60 (52,68)
30-day rebleeding	Cutoff value	12	1	3	5
	AUC (95% CI)	0·63 (0.54,0.72)	0.60 (0.52,0.69)	0.59 (0.50,0.68)	0.63 (0.54,0.71)
	Sensitivity† (95% CI)	78 (64,88)	92 (80,97)	74 (60,85)	72 (58,84)
	Specificity† (95% CI)	44 (34,53)	34 (25,43)	44 (33,52)	53 (42,61)
	PPV† (95% CI)	39 (34,44)	38 (35,42)	38 (32,43)	41 (35,47)
	NPV† (95% CI)	81 (71,88)	90 (77,96)	78 (58,85)	80 (71,86)
	PLR (95% CI)	1·40 (1.12,1.75)	1.39 (1.19,1.63)	1.31 (1.04,1.65)	1.52 (1.17,1.97)
	NLR (95% CI)	0·49 (0.28,0.86)	0.24 (0.09,0.63)	0.59 (0.35,0.99)	0.52 (0.32,0.85)
	Accuracy† (95% CI)	55 (46,62)	52 (44,60)	53 (45,61)	58 (50,66)
Composite outcome	Cutoff value	12	2	3	3
	AUC (95% CI)	0.75 (0.66,0.85)	0.76 (0.68,0.85)	0.69 (0.58,0.81)	0.72 (0.61,0.83)
	Sensitivity† (95% CI)	71 (62,79)	60 (51,68)	70 (61,77)	93 (87,96)
	Specificity† (95% CI)	73 (54,87)	87 (69,96)	70 (50,85)	47 (28,65)
	PPV† (95% CI)	92 (86,95)	95 (88,98)	91 (85,94)	88 (84,91)
	NPV† (95% CI)	37 (29,45)	33 (28,39)	35 (27,43)	60 (42,76)
	PLR (95% CI)	2·68 (1.47,4.90)	4.53(1.80,11.41)	2.33 (1.34,4.08)	1.75 (1.24,2.45)
	NLR (95% CI)	0.39 (0.27,0.55)	0.46 (0.35,0.59)	0.43 (0.30,0.61)	0.15 (0.07,0.31)
	Accuracy† (95% CI)	72 (64,78)	65 (57,72)	70 (62,76)	84 (77,89)

AUC: Area under curves; CI: Confidence interval; IQR: Interquartile range; PPV: Positive predictive value; NPV: Negative predictive value; PLR: Positive likelihood ratio; NLR: Negative likelihood ratio

* The cutoff values of each test were selected separately for each outcome using Youden's index

† All the sensitivity, specificity, PPV, NPV, and accuracy were presented as percentages

significant practical use in either variceal or non-variceal UGIB in several studies (27). Considering the differences

in pathophysiology and management between variceal and non-variceal UGIB, several studies have specifically excluded variceal UGIB patients. This approach allows for a focused comparison of the efficacy of scoring systems in the more prevalent non-variceal UGIB group. By excluding variceal UGIB cases, these studies aim to provide a clearer understanding of the performance and applicability of scoring systems in the context of non-variceal UGIB, which is more common in clinical practice (31-34).

In our study, the AUC values obtained for the evaluated scoring systems were generally consistent with previous research for the respective outcomes. However, we observed lower accuracy for the GBS compared to some of the previous studies. It is worth noting that, apart from AIMS65 for predicting in-hospital mortality, most of the AUC values fell within the range of 0.6 to 0.7, which may not be considered satisfactory for practical management of a high-risk population. These findings suggest that, while scoring systems can provide useful adjunctive measures, they should not be solely relied upon. Instead, frequent reassessments and monitoring of the patients' clinical course should be integrated into the management approach. This emphasizes the importance of a comprehensive and multifaceted approach to the practical management of high-risk patients.

5. Limitations

Our study had several limitations that should be acknowledged. One significant limitation is the high mortality rate observed in our study population compared to other reports. This higher mortality rate can be attributed to our exclusion criteria, which involved excluding patients with probable causes of death related to other accompanying illnesses. Additionally, the exclusion of variceal bleeding cases and patients who did not undergo endoscopy could have contributed to a selection bias, potentially impacting the overall mortality rate observed in our study. Lastly, a relatively small sample size and data limited to a single center are limitations that originate from our limited resources.

6. Conclusion

AIMS65 may be the most powerful scoring system for prediction of in-hospital mortality while also displaying considerable efficacy in anticipating the need for massive transfusions. Additionally, GBS and AIMS65 could be moderately and cautiously relied on for preparations regarding the need for PRBC transfusion. The other associations and predictions were not in a range that could be recommended for practical purposes according to our results, and overreliance on a specific scoring system without concern for implicit clinical judgment and trends of signs or symptoms might lead to erroneous predictions and disastrous outcomes.

7. Declarations

7.1. Acknowledgement

None.

7.2. Authors' contribution

Conceptualization: NT, MM, AM; Design: MM, AM; Supervision: PP, AM; Resources: MRZR, HM, SDA; Materials: NT, AM; Data collection and/or processing: MRZR, HM, MJN, AH; Analysis and/or interpretation: MRZR, MI; Literature search: MRZR, AM, HM; Writing manuscript: MRZR, HM, AM, SDA; Critical review: PP, MM.

7.3. Conflict of interest

All authors have no potential conflicts of interest.

7.4. Funding

This research received no funding.

7.5. Data Availability Statement

The data that supports the findings of this study is available from the corresponding author upon reasonable request.

References

1. Kamboj AK, Hoversten P, Leggett CL. Upper gastrointestinal bleeding: etiologies and management. Mayo Clin Proc. 2019;94(4):697-703.
2. Stanley AJ, Laine L. Management of acute upper gastrointestinal bleeding. BMJ. 2019 ;364:l536.
3. Rajan SS, Sawe HR, Iyullu AJ, Kaale DA, Olambo NA, Mfinanga JA, et al. Profile and outcome of patients with upper gastrointestinal bleeding presenting to urban emergency departments of tertiary hospitals in Tanzania. BMC Gastroenterol. 2019;19:212.
4. Vora P, Pietila A, Peltonen M, Brobert G, Salomaa V. Thirty-year incidence and mortality trends in upper and lower gastrointestinal bleeding in Finland. JAMA Netw Open. 2020;3(10):e2020172.
5. Van Leerdam ME, Vreeburg EM, Rauws EAJ, Geraedts AAM, Tijssen JGP, Reitsma JB, et al. Acute upper GI bleeding: did anything change? Time trend analysis of incidence and outcome of acute upper GI bleeding between 1993/1994 and 2000. Am J Gastroenterol. 2003;98(7):1494-9.
6. Abougergi MS, Travis AC, Saltzman JR. The in-hospital mortality rate for upper GI hemorrhage has decreased over 2 decades in the United States: a nationwide analysis. Gastrointest Endosc. 2015;81(4):882-8.e1.
7. Wuerth BA, Rockey DC. Changing epidemiology of upper gastrointestinal hemorrhage in the last decade: a nationwide analysis. Dig Dis Sci. 2018;63:1286-93.
8. Laursen SB. Treatment and prognosis in peptic ulcer bleeding. Dan Med J. 2014;61(1):B4797.

9. Monteiro S, Gonçalves TC, Magalhães J, Cotter J. Upper gastrointestinal bleeding risk scores: who, when and why? World J Gastrointest Pathophysiol. 2016;7(1):86.
10. Bakhtavar HE, Bagi HRM, Rahmani F, Ettehadi A. Clinical scoring systems in predicting the outcome of acute upper gastrointestinal bleeding; a narrative review. Emergency. 2017;5(1).
11. Custovic N, Husic-Selimovic A, Srsen N, Prohic D. Comparison of Glasgow-Blatchford score and Rockall score in patients with upper gastrointestinal bleeding. Med Arch. 2020;74(4):270.
12. Oakland K. Risk stratification in upper and upper and lower GI bleeding: Which scores should we use? Best Pract Res Clin Gastroenterol. 2019;42:101613.
13. Barkun AN, Almadi M, Kuipers EJ, Laine L, Sung J, Tse F, et al. Management of nonvariceal upper gastrointestinal bleeding: guideline recommendations from the international consensus group. Ann Intern Med. 2019;171(11):805-22.
14. Sung JJ, Chiu PW, Chan FKL, Lau JY, Goh KL, Ho LH, et al. Asia-Pacific working group consensus on non-variceal upper gastrointestinal bleeding: an update 2018. Gut. 2018;67(10):1757-68.
15. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for uppergastrointestinal haemorrhage. The Lancet. 2000;356(9238):1318-21.
16. Saltzman JR, Tabak YP, Hyett BH, Sun X, Travis AC, Johannes RS. A simple risk score accurately predicts in-hospital mortality, length of stay, and cost in acute upper GI bleeding. Gastrointest Endosc. 2011;74(6):1215-24.
17. Rockall TA, Logan RF, Devlin HB, Northfield TC. Risk assessment after acute upper gastrointestinal haemorrhage. Gut. 1996;38(3):316-21.
18. Gu L, Xu F, Yuan J. Comparison of AIMS65, Glasgow–Blatchford and Rockall scoring approaches in predicting the risk of in-hospital death among emergency hospitalized patients with upper gastrointestinal bleeding: a retrospective observational study in Nanjing, China. BMC Gastroenterol. 2018;18(1):1-8.
19. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden index and optimal cut-point estimated from observations affected by a lower limit of detection. Biom J. 2008;50(3):419-30.
20. Deeks J, Altman DG. Diagnostic tests 4: likelihood ratios. Bmj. 2004;329(7458):168-9.
21. Mokhtare M, Bozorgi V, Agah S, Nikkhah M, Faghihi A, Boghratian A, et al. Comparison of Glasgow-Blatchford score and full Rockall score systems to predict clinical outcomes in patients with upper gastrointestinal bleeding. Clin Exp Gastroenterol. 2016;337-43.
22. Chandra S, Hess EP, Agarwal D, Nestler DM, Montori VM, Song LMWK, et al. External validation of the Glasgow-Blatchford bleeding score and the Rockall score in the US setting. Am J Emerg Med. 2012 ;30(5):673-9.
23. Bryant RV, Kuo P, Williamson K, Yam C, Schoeman MN, Holloway RH, et al. Performance of the Glasgow-Blatchford score in predicting clinical outcomes and intervention in hospitalized patients with upper GI bleeding. Gastrointest Endosc. 2013;78(4):576-83.
24. Robertson M, Majumdar A, Boyapati R, Chung W, Worland T, Terbah R, et al. Risk stratification in acute upper GI bleeding: comparison of the AIMS65 score with the Glasgow-Blatchford and Rockall scoring systems. Gastrointest Endosc. 2016;83(6):1151-60.
25. Martínez-Cara JG, Jiménez-Rosales R, Úbeda-Muñoz M, de Hierro ML, de Teresa J, Redondo-Cerezo E. Comparison of AIMS65, Glasgow–Blatchford score, and Rockall score in a European series of patients with upper gastrointestinal bleeding: performance when predicting in-hospital and delayed mortality. United Eur Gastroenterol J. 2016;4(3):371-9.
26. Totagi A, Srinivas U, Paramasivan P, Krishnan S, Palaniswamy K, Mohan A, et al. Comparison of complete Rockall score, Glasgow–Blatchford score, and AIMS 65 score for predicting in-hospital mortality in patients presenting with upper gastrointestinal hemorrhage at a tertiary care hospital in southern India. Gastroenterol Hepatol Endosc Pract. 2021;1(2):55.
27. Chang A, Ouejiaraphant C, Akarapatima K, Rattanasupa A, Prachayakul V. Prospective comparison of the AIMS65 score, Glasgow-Blatchford score, and Rockall score for predicting clinical outcomes in patients with variceal and nonvariceal upper gastrointestinal bleeding. Clin Endosc. 2021;54(2):211-21. 2020;4(4):e89.
28. Tang Y, Shen J, Zhang F, Zhou X, Tang Z, You T. Scoring systems used to predict mortality in patients with acute upper gastrointestinal bleeding in the ED. Am J Emerg Med. 201;36(1):27-32.
29. Park SM, Yeum SC, Kim BW, Kim JS, Kim JH, Sim EH, et al. Comparison of AIMS65 score and other scoring systems for predicting clinical outcomes in Koreans with nonvariceal upper gastrointestinal bleeding. Gut Liver. 2016;10(4):526-31.
30. Lai YC, Hung MS, Chen YH, Chen YC. Comparing AIMS65 score with MEWS, qSOFA score, Glasgow-Blatchford score, and Rockall score for predicting clinical outcomes in cirrhotic patients with upper gastrointestinal bleeding. J Acute Med. 201;8(4):154-67.
31. Marmo R, Koch M, Cipolletta L, Capurso L, Grossi E, Cestari R, et al. Predicting mortality in non-variceal upper gastrointestinal bleeders: validation of the Italian PNED score and prospective comparison with the Rockall score. Am J Gastroenterol. 2010;105(6):1284-91.
32. Kim MS, Choi J, Shin WC. AIMS65 scoring system is comparable to Glasgow-Blatchford score or Rockall score for prediction of clinical outcomes for nonvariceal upper gastrointestinal bleeding. BMC Gastroenterol. 2019;19(1):136.
33. Yang HM, Jeon SW, Jung JT, Lee DW, Ha CY, Park KS, et al. Comparison of scoring systems for nonvariceal upper

gastrointestinal bleeding: a multicenter prospective cohort study. J Gastroenterol Hepatol. 2016;31(1):119-25.
34. Tandon P, Bishay K, Fisher S, Yelle D, Carrigan I, Wooller K, et al. Comparison of clinical outcomes between variceal and non-variceal gastrointestinal bleeding in patients with cirrhosis. J Gastroenterol Hepatol. 2018;33(10):1773-9.