



Comparison of risk scoring systems for patients presenting with upper gastrointestinal bleeding: international multicentre prospective study

Adrian J Stanley,¹ Loren Laine,² Harry R Dalton,³ Jing H Ngu,⁴ Michael Schultz,^{5,7} Roseta Abazi,⁶ Liam Zakko,² Susan Thornton,¹ Kelly Wilkinson,³ Christopher J L Khor,⁴ Iain A Murray,³ Stig B Laursen⁶ on behalf of the International Gastrointestinal Bleeding Consortium

¹Department of Gastroenterology, Glasgow Royal Infirmary, Glasgow G4 OSF, UK

²Section of Digestive Diseases, Yale School of Medicine, New Haven, and VA Connecticut Healthcare System, West Haven, CT, USA

³Gastrointestinal Unit, Royal Cornwall Hospital, Cornwall, UK

⁴Department of Gastroenterology and Hepatology, Singapore General Hospital, Singapore

⁵Department of Medicine, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

⁶Department of Medical Gastroenterology, Odense University Hospital, Odense, Denmark

⁷Gastroenterology Unit, Southern District Health Board, Dunedin Hospital, Dunedin, New Zealand

Correspondence to: A J Stanley adrian.stanley@ggc.scot.nhs.uk
Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2017;356:i6432
<http://dx.doi.org/10.1136/bmj.i6432>

Accepted: 25 November 2016

ABSTRACT

OBJECTIVE

To compare the predictive accuracy and clinical utility of five risk scoring systems in the assessment of patients with upper gastrointestinal bleeding.

DESIGN

International multicentre prospective study.

SETTING

Six large hospitals in Europe, North America, Asia, and Oceania.

PARTICIPANTS

3012 consecutive patients presenting over 12 months with upper gastrointestinal bleeding.

MAIN OUTCOME MEASURES

Comparison of pre-endoscopy scores (admission Rockall, AIMS65, and Glasgow Blatchford) and post-endoscopy scores (full Rockall and PNED) for their ability to predict predefined clinical endpoints: a composite endpoint (transfusion, endoscopic treatment, interventional radiology, surgery, or 30 day mortality), endoscopic treatment, 30 day mortality, rebleeding, and length of hospital stay. Optimum score thresholds to identify low risk and high risk patients were determined.

RESULTS

The Glasgow Blatchford score was best (area under the receiver operating characteristic curve (AUROC) 0.86) at predicting intervention or death compared with the full Rockall score (0.70), PNED score (0.69), admission

Rockall score (0.66, and AIMS65 score (0.68) (all $P < 0.001$). A Glasgow Blatchford score of ≤ 1 was the optimum threshold to predict survival without intervention (sensitivity 98.6%, specificity 34.6%). The Glasgow Blatchford score was better at predicting endoscopic treatment (AUROC 0.75) than the AIMS65 (0.62) and admission Rockall scores (0.61) (both $P < 0.001$). A Glasgow Blatchford score of ≥ 7 was the optimum threshold to predict endoscopic treatment (sensitivity 80%, specificity 57%). The PNED (AUROC 0.77) and AIMS65 scores (0.77) were best at predicting mortality, with both superior to admission Rockall score (0.72) and Glasgow Blatchford score (0.64; $P < 0.001$). Score thresholds of ≥ 4 for PNED, ≥ 2 for AIMS65, ≥ 4 for admission Rockall, and ≥ 5 for full Rockall were optimal at predicting death, with sensitivities of 65.8-78.6% and specificities of 65.0-65.3%. No score was helpful at predicting rebleeding or length of stay.

CONCLUSIONS

The Glasgow Blatchford score has high accuracy at predicting need for hospital based intervention or death. Scores of ≤ 1 appear the optimum threshold for directing patients to outpatient management. AUROCs of scores for the other endpoints are less than 0.80, therefore their clinical utility for these outcomes seems to be limited.

TRIAL REGISTRATION

Current Controlled Trials ISRCTN16235737.

Introduction

Upper gastrointestinal bleeding is a common cause of admission to hospital worldwide, with a UK incidence of 103-172 per 100 000 adults each year and mortality of 8-14%.¹⁻³ Many risk assessment scores have been developed to predict clinically relevant outcomes, including mortality, need for hospital based intervention, rebleeding, and length of hospital stay.⁴⁻⁹ Several of these, including the Rockall score and recently described progetto nazionale emorragia digestiva (PNED) score require endoscopy before calculation.^{4,5} The Rockall score has been widely reported and it has been suggested that the PNED score is superior to it, although it has not been externally validated.⁵ Other endoscopy based scores have been described but are not appropriate for unselected patients with upper gastrointestinal bleeding, have not been externally validated, or have been shown to be inferior to the Rockall score or PNED score.¹⁰ However, requiring endoscopy to calculate a score might delay risk assessment in

WHAT IS ALREADY KNOWN ON THIS TOPIC

Several pre-endoscopy and post-endoscopy risk scores have been reported to predict a variety of endpoints in patients presenting with upper gastrointestinal bleeding

Using low and high scores to direct patient care to outpatient management or urgent endoscopy and higher level care has been suggested

The accuracy and generalisability of these scores and the optimum thresholds to predict low risk and high risk patients to help direct care remain unclear

WHAT THIS STUDY ADDS

This international study suggests that the Glasgow Blatchford score is accurate in predicting need for intervention, or death at all centres, with a score ≤ 1 the optimal threshold for directing low risk patients to outpatient management

All scores have low predictive accuracy for other specific outcomes, including endoscopic treatment and mortality, therefore their clinical utility to direct management of high risk patients seems limited

Widespread use of the Glasgow Blatchford score in the emergency departments of many countries might allow up to 19% of patients with upper gastrointestinal bleeding to be safely managed as outpatients

some healthcare settings, as there can be considerable delays in performing endoscopy out of hours or on weekends.¹¹

Recently, much interest has been shown in pre-endoscopic risk scores for upper gastrointestinal bleeding, which can be calculated shortly after presentation to hospital. The most widely studied scores are the abbreviated “admission” Rockall score, Glasgow Blatchford score, and recently described AIMS65 (albumin level <30 g/L (A), international normalised ratio >1.5 (I), altered mental status (M), systolic blood pressure ≤90 mm Hg (S), and age >65 years (65)) score.^{7,8} These scoring systems use clinical, haemodynamic, and (for Glasgow Blatchford and AIMS65 scores) readily available laboratory variables. Some studies have suggested that these scores could be used to identify patients at very low risk who could be managed as outpatients.¹²⁻¹⁴ Studies have also suggested that these scores could identify patients at higher risk who might require urgent endoscopy or management in high dependency or intensive care units.^{15,16}

Several studies have compared scores in their ability to predict various outcomes.^{12,16-19} Although the Glasgow Blatchford score seems useful in identifying very low risk patients in the UK many studies have been small and single centre and no international study has compared all the commonly used scores. Recent guidelines have suggested use of risk scores for patients with upper gastrointestinal bleeding, but uncertainty remains about their exact role in clinical practice.²⁰⁻²³ In this prospective international study we compared five endoscopic and pre-endoscopic risk assessment scores for their ability to predict clinically relevant endpoints. In addition we assessed the clinical utility of these scores, by determining optimal thresholds for identifying patients at very low risk who could be managed as outpatients, and higher risk patients who might require specific management strategies aimed at improving outcome.

Methods

The study was undertaken in Yale-New Haven Hospital (USA), Glasgow Royal Infirmary (Scotland), Royal Cornwall Hospital Truro (England), Odense University Hospital (Denmark), Singapore General Hospital (Singapore), and Dunedin Hospital (New Zealand).

The predetermined clinical endpoints were the composite endpoint of need for hospital based intervention (red blood cell transfusion, endoscopic treatment, interventional radiology, or surgery) or death, endoscopic treatment, 30 day mortality, rebleeding within seven days, and length of hospital stay.

Patient management

Patients were included if they presented to the hospital with evidence of upper gastrointestinal bleeding defined by haematemesis, coffee-ground vomiting, or melaena. We excluded patients who developed upper gastrointestinal bleeding while an inpatient for another reason.

All patients presenting to each hospital with upper gastrointestinal bleeding were initially assessed in the

emergency department or acute assessment unit. Two centres (Glasgow and Odense) had a policy of non-admission for those with Glasgow Blatchford scores of 1 or less, and one centre (Truro) for those with Glasgow Blatchford scores of 2 or less and aged less than 70 years, unless required for other reasons. Outpatient endoscopy was arranged in all three centres. Proton pump inhibitors (PPIs) were not routinely given to all admitted patients before endoscopy. The practice in all centres was to perform endoscopy within 24 hours in all admitted patients where possible. After endoscopy, the policy in all centres was to administer high dose PPIs by intravenous bolus followed by infusion to patients with high risk ulcer stigmata who required endoscopic treatment, and to other selected patients depending on clinical judgment.²⁰⁻²³ For patients with suspected variceal bleeding, the policy in all centres was to give intravenous vasopressors and antibiotics before endoscopy.^{21,24}

The endoscopic practice in all centres for patients with high risk stigmata of non-variceal bleeding was by injection of dilute adrenaline (epinephrine) into and around the bleeding point, thermal contact or clips, or both, but not adrenaline alone.²⁰⁻²³ Band ligation or injection of tissue glue with or without transjugular intrahepatic portosystemic shunt was performed in cases of oesophageal or gastric variceal bleeding, respectively.^{21,24} In line with recent evidence and guidelines, the policy was to administer red blood cells at a haemoglobin threshold of 70-80 g/L (7-8 g/dL), or as guided by the clinician in patients with severe haemorrhage.^{24,25}

Data collection

At each centre, data on consecutive, unselected patients presenting to the hospital with upper gastrointestinal bleeding were collected over a 12 month period between March 2014 and March 2015. Data on patients managed as outpatients in Glasgow, Truro, and Odense were also included. A dedicated research nurse, doctor, or medical student collected data at each site. The data collected included patient characteristics and haemodynamic and laboratory variables at presentation necessary to calculate the full Rockall, admission Rockall, Glasgow Blatchford, AIMS65, and PNED scores (see supplementary table 1). Endoscopic findings were recorded.

Blood transfusion, endoscopic treatment, and interventional radiology or surgery were recorded, as was rebleeding within seven days, as previously defined (see supplementary table 2),²⁶ 30 day mortality, and length of hospital stay. Supplementary table 3 shows the collected data.

Statistical analysis

We compared each scores' ability to predict the predetermined outcomes, using calculation of area under the receiver operating characteristic curves (AUROCs) and 95% confidence intervals. All comparisons were based on patients in whom all the compared scores could be calculated. The AUROCs were considered dependent

and were compared based on the method by Delong et al.²⁷ The optimal score thresholds to predict patients at very low risk who are potentially suitable for outpatient management were identified based on a sensitivity of 95% or more. For the other outcomes, we considered thresholds associated with maximum Youden index²⁸ to be optimal. In addition, the performance of the scores was assessed by calculation of sensitivity, specificity, positive predictive value, negative predictive value, and proportion of patients classified as low risk or high risk. We calculated rates of transfusion, endoscopic treatment, and mortality for those classified as low risk. Proportions were compared using Pearson's χ^2 test and Fischer's exact test. Medians were compared using the Mann-Whitney U test and the Kruskal-Wallis equality of populations rank test.

We aimed for an adequate sample size to compare all the predetermined outcomes. Existing data indicated that the highest power was needed for comparative analyses on mortality. Therefore assuming AUROCs of 0.86 for mortality, 0.78 for AIMS65 score^{8,20} and full Rockall score,^{10,21} with an α of 5%, power of 90%, rank correlation of 0.5, and mortality of 5%,¹³ we determined that we required a sample size of 2814 patients. We used a two tailed significance level of 5%. Data analysis was undertaken using STATA 11.0 (StataCorp, College Station, TX).

Only the local principal investigator and dedicated nurse, junior doctor, or medical student had access to the identifiable data from their centre, with the anonymised data sent to SBL and AJS for central analysis. Our report follows STROBE guidelines.

Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for recruitment, design, or implementation of the study. No patients were asked to advise on interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

Results

Patient characteristics and baseline scores

A total of 3012 patients were included in the study, with 2868 (95%) followed up for 30 days. The median age of patients was 65 years, and 58% were men. Table 1 shows the patients' characteristics, endoscopic findings, interventions, and outcomes. A total of 1348 (45%) patients needed hospital based intervention or died within 30 days. Endoscopic treatment was performed in 574 (19%), 37 (1.2%) required interventional radiology or surgery, and 144 (5%) rebled within seven days. The median length of stay was 3 days (95% confidence interval 0 to 16 days), and 30 day mortality was 7%. Overall, the mean Glasgow Blatchford score was 6.6 (95% confidence interval 0 to 14), mean AIMS65 score was 1 (0 to 3), mean admission Rockall score was 2.7 (0 to 5), mean full Rockall score was 3.9 (1 to 7), and mean PNED score was 3.0 (0 to 8). Supplementary table 4 show these data by centre.

Comparison of scores' ability to predict outcomes Intervention or mortality

The Glasgow Blatchford score had the highest discriminative ability (AUROC 0.86) at predicting need for intervention or death compared with the PNED score (0.69; $P<0.001$), full Rockall score (0.70; $P<0.001$), admission Rockall score (0.66 $P<0.001$), and AIMS65 score (0.68; $P<0.001$) (table 2 and fig 1). The Glasgow Blatchford score performed consistently well in all centres, with AUROCs ranging between 0.85 and 0.91 (see supplementary table 5). However, performance of AIMS65 and the full Rockall scoring systems varied by centre, with AUROCs ranging from 0.54-0.75 and 0.57-0.79, respectively. AUROCs among centres were 0.59-0.72 for admission Rockall score and 0.70-0.76 for PNED score (see supplementary table 5).

Tables 3 shows the ability of the three pre-endoscopic scores to predict patients at very low risk who did not require intervention and survived at optimal cut-offs. A Glasgow Blatchford score of 1 or less had a sensitivity of 98.6%, specificity of 34.6%, positive predictive value of 96.6%, and negative predictive value of 56.0% for this combined endpoint. Intervention or death was recorded in 3.4% of the 564 patients with Glasgow Blatchford scores of 1 or less, compared with 14% of the

Table 1 | Characteristics, treatment, and outcome of patients (n=3012). Values are numbers (percentages) unless stated otherwise

Characteristics	Data
Median (95% CI) age (years)	65 (24 to 90)
Men	1750 (58)
Comorbidity:	
Ischaemic heart disease	580 (19)
Liver disease	453 (15)
Renal failure	266 (9)
Any malignancy	430 (14)
Median (95% CI) systolic blood pressure (mm Hg)	125 (90 to 170)
Median (95% CI) pulse	89 (61 to 126)
Median (95% CI) haemoglobin level (g/L)	112 (58 to 162)
Findings at endoscopy:	
Nothing abnormal	297 (14)
Erosive disease*	583 (28)
Gastric/duodenal ulcer	572 (28)
Variceal bleeding	143 (7)
Upper gastrointestinal cancer	70 (3)
No endoscopy	937 (31)
Median (95% CI) time to endoscopy (hours)	19 (3 to 72)
Treatment:	
Median (95% CI) No of transfusions	1.3 (0 to 6)
Endoscopic treatment	574 (19)
Surgery/interventional radiology	37 (1.2)
Outcome:	
No need for intervention, or death	1636 (55)
Rebleeding	144 (5)
Mortality	208 (7)
Mean (95% CI) score:	
Glasgow Blatchford	6.6 (0 to 14)
AIMS65	1.0 (0 to 3)
Admission Rockall	2.7 (0 to 5)
Full Rockall	3.9 (1 to 7)
PNED	3.0 (0 to 8)

*Oesophagitis, gastritis, or duodenitis.

Outcome by scoring system	AUROC (95% CI)
Intervention or death:	
Glasgow Blatchford	0.89 (0.87 to 0.90)
AIMS65	0.70 (0.68 to 0.72)
Admission Rockall	0.69 (0.67 to 0.71)
Full Rockall	0.69 (0.67 to 0.71)
PNED	0.71 (0.70 to 0.73)
Need for endoscopic treatment:	
Glasgow Blatchford	0.75 (0.73 to 0.77)
AIMS65	0.63 (0.60 to 0.65)
Admission Rockall	0.61 (0.59 to 0.64)
Rebleeding:	
Glasgow Blatchford	0.70 (0.66 to 0.74)
AIMS65	0.62 (0.57 to 0.66)
Admission Rockall	0.62 (0.57 to 0.66)
Full Rockall	0.63 (0.58 to 0.68)
PNED	0.85 (0.83 to 0.88)
Mortality:	
Glasgow Blatchford	0.69 (0.66 to 0.72)
AIMS65	0.78 (0.75 to 0.81)
Admission Rockall	0.76 (0.73 to 0.79)
Full Rockall	0.72 (0.68 to 0.77)
PNED	0.79 (0.76 to 0.82)

AUROC=Area under receiver operating characteristic curve.
Values represent all patients with available data.

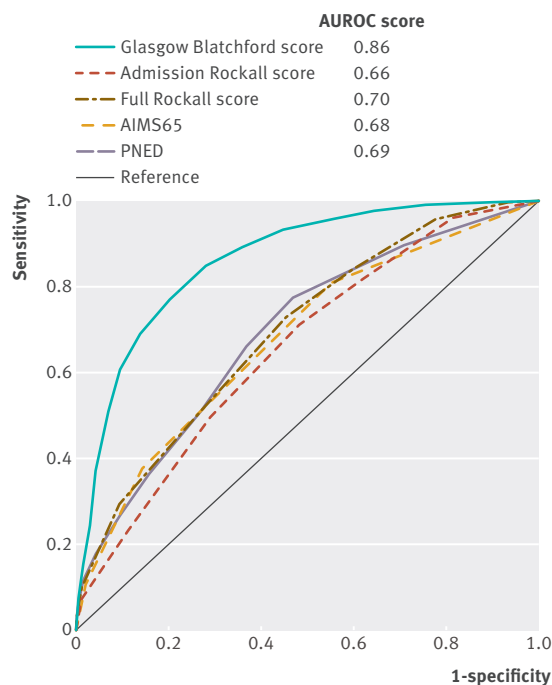


Fig 1 | Comparisons of scores in prediction of need for any intervention (transfusion, endoscopic treatment, interventional radiology or surgery) or 30 day mortality (n=1704). All figures compare patients with complete data for all compared scores. AUROC=area under the receiver operating characteristic curve

436 patients with admission Rockall scores of 0, and 25% of the 865 patients with AIMS65 scores of 0 (see table 3). The most common intervention in those with Glasgow Blatchford scores of 1 or less was blood transfusion, with endoscopic treatment performed in 1.4%

and overall mortality 0.4%. The proportion of low risk patients defined by Glasgow Blatchford scores of 1 or less varied by site, ranging from 9% in Dunedin to 31% in Glasgow.

Endoscopic treatment

The Glasgow Blatchford score (AUROC 0.75) performed better in predicting need for endoscopic treatment than the other pre-endoscopic scores: AIMS65 (0.62) and admission Rockall (0.61; $P<0.001$ for both; fig 2). The Glasgow Blatchford scoring system performed equally well in predicting this endpoint in all centres (see supplementary table 5). A threshold Glasgow Blatchford score of 7 or more was best at predicting endoscopic treatment, with a sensitivity of 80.4%, specificity of 57.4%, positive predictive value of 31.3%, and negative predictive value of 92.4% (table 4).

Mortality

The AIMS65 score (AUROC 0.77) and PNED score (0.77) had similar discriminative abilities for predicting 30 day mortality (table 2 and fig 3). The PNED score was better at predicting mortality than the admission Rockall score (0.72; $P=0.05$), full Rockall score (0.72; $P=0.05$), and Glasgow Blatchford score (0.64; $P<0.001$). AIMS65 was better at predicting mortality than the Glasgow Blatchford score ($P<0.001$) and admission Rockall score ($P=0.05$). Furthermore, the AIMS65 score had a near statistically significantly higher AUROC compared with the full Rockall score ($P=0.06$). The best score thresholds at predicting 30 day mortality were 4 or more for PNED, 2 or more for AIMS65, 4 or more for admission Rockall, 5 or more for full Rockall, and 5 or more for Glasgow Blatchford. Table 5 shows the predictive ability of all five scores to identify high risk patients.

Rebleeding and length of stay

PNED, which includes rebleeding in the score, was best (AUROC 0.85) at predicting rebleeding compared with the Glasgow Blatchford score (0.66; $P<0.001$), full Rockall score (0.64; $P<0.001$), admission Rockall score (0.62; $P<0.001$), and AIMS65 (0.60; $P<0.001$). The Glasgow Blatchford score was better at predicting rebleeding compared with the AIMS65 score ($P=0.04$). All risk scores had poor discriminative abilities for predicting length of stay for more than three days. The AUROCs were 0.68 for PNED score, 0.65 for full Rockall score, 0.64 for admission Rockall score, 0.63 for Glasgow Blatchford score, and 0.61 for AIMS65 score.

Missing values

Values were missing for the Glasgow Blatchford score (n=80), AIMS65 score (n=511), admission Rockall score (n=43), full Rockall score (n=1000), PNED score (n=178), comorbidity (n=1), systolic blood pressure (n=41), pulse (n=38), haemoglobin concentration (n=28), findings at endoscopy (n=2), true low risk status (n=28), total units of blood transfused (n=23), performance of endoscopic treatment (n=20), performance of surgery or interventional radiology (n=5), rebleeding (n=51), and mortality (n=1). Comparisons of AUROCs were undertaken using

Table 3 | Outcomes at optimal thresholds of Glasgow Blatchford score, AIMS65, and admission Rockall score in prediction of need for intervention, or death. Values are numbers (percentages) unless stated otherwise

Scoring system	Cut-off	Low risk patients*	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Received transfusion	Endoscopic treatment	Surgery or interventional radiology†	Mortality‡
Glasgow Blatchford	≤1	564 (19.2)	98.6	34.6	96.6	56.0	10 (1.8)	8 (1.4)	2 (0.4)	2 (0.4)
AIMS65	0	865 (34.6)	81.5	49.9	74.7	59.9	165 (19)	107 (12)	9 (1.0)	6 (0.7)
Admission Rockall	0	436 (14.7)	95.6	23.4	86.5	50.9	41 (9.4)	29 (6.7)	0 (0)	1 (0.2)

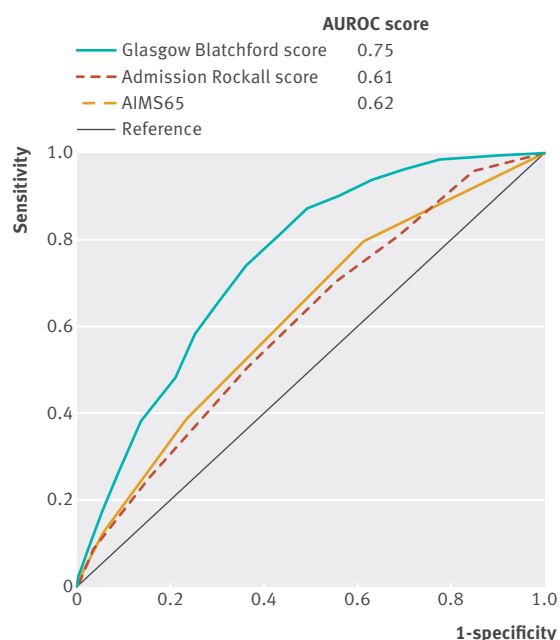
PPV=positive predictive value; NPV=negative predictive value.

Several patients fulfilled more than one endpoint.

*Classified as low risk according to risk scoring system.

†Number (%) of patients needing surgery, or arterial embolisation among patients classified as low risk.

‡Number (%) of patients dying within 30 days from presentation among patients classified as low risk.

**Fig 2 | Comparisons of pre-endoscopic scores in prediction of need for endoscopic treatment (n=2478). AUROC=area under the receiver operating characteristic curve**

complete case analysis. Data for all pre-endoscopic scores (admission Rockall, AIMS65, Glasgow Blatchford) were available in 2495 (83%) patients, and data for both post-endoscopy scores (full Rockall, PNED) in 1922 (64%). Data for the endoscopic treatment endpoint and the composite endpoint were available in 2478 (82%) and 1707 (57%) patients, respectively.

Patients with missing values for full Rockall or AIMS65 had lower Glasgow Blatchford scores (mean: 4.8 (95% confidence interval 0.0 to 13.0) v 7.8 (1 to 14); $P<0.001$), PNED scores (mean: 2.6 (0.0 to 7.0) v 3.2 (0.0 to 8.0); $P<0.001$), and admission Rockall scores (mean: 2.5

(0 to 5) v 2.9 (0 to 5); $P<0.001$) compared with patients without missing data for these two scores. Patients with missing values for full Rockall score or AIMS65 less often underwent hospital based intervention or died (27% v 58%; $P<0.001$) but had similar 30 day mortality (7.5% v 6.5%; $P=0.32$) compared with patients without missing values for these scores. Analysis from the centre with the lowest missing data on AIMS65 and full Rockall scores (Odense, at 1.9% and 8.3% respectively) showed that the Glasgow Blatchford score had a higher discriminative ability to predict need for intervention, or death, and need for endoscopic treatment, compared with all other risk scores ($P<0.001$; see supplementary table 5).

Discussion

This international multicentre study shows that the Glasgow Blatchford score is an accurate risk score at all sites for predicting need for clinical intervention, or death, after upper gastrointestinal bleeding. Although scores for AIMS65 and PNED are the best for predicting mortality and the Glasgow Blatchford score is best for predicting endoscopic treatment, accuracy and clinical utility is relatively low for these endpoints.

Strengths and limitations of this study

Strengths of our study include its international multicentre design and large number of consecutive patients studied. By assessing the five pre-endoscopy and post-endoscopy scores that appeared most promising for clinical use, we could investigate the optimum way to risk assess patients with upper gastrointestinal bleeding, both early after presentation and after endoscopic diagnosis and treatment. We believe our choice of several clinically relevant endpoints in addition to a composite endpoint to compare the scores provides clinicians with useful information. Clinicians may be keen to identify those at high risk for need of endoscopic

Table 4 | Need for endoscopic treatment at optimal score thresholds of each of the pre-endoscopic scores: Glasgow Blatchford score, AIMS65, and admission Rockall score

Scoring system	Cut-off	No (%) of patients classified as high risk	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Glasgow Blatchford	≥7	1456 (50)	80.4	57.4	31.3	92.4
AIMS65	≥1	1619 (65)	79.7	38.7	25.9	87.6
Admission Rockall	≥3	1686 (57)	69.8	45.9	23.5	86.5

PPV=positive predictive value; NPV=negative predictive value.

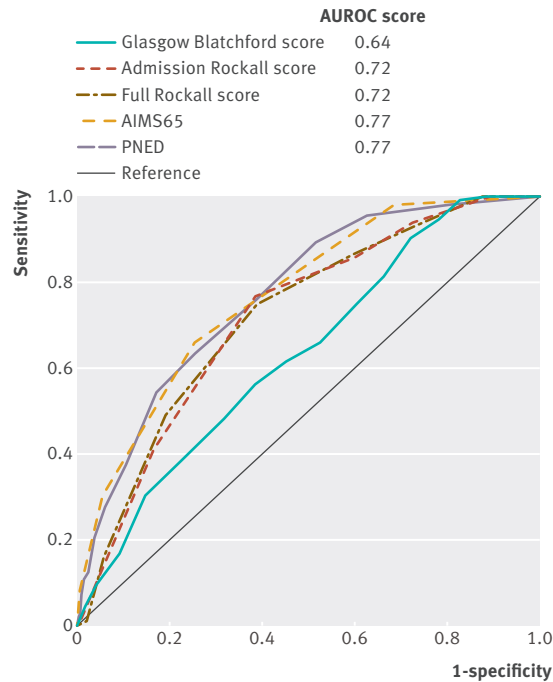


Fig 3 | Comparisons of scores in prediction of 30 day mortality (n=1707). AUROC=area under the receiver operating characteristic curve

treatment, to direct such patients to urgent endoscopy. Similarly, those at high risk of death could be moved to areas of higher level care. While a potential weakness of composite endpoints is that they may be driven largely by one or two components and thus might not provide useful information if the components are not clinically appropriate to combine, we believe our composite endpoint is clinically useful because it includes the outcomes that generally would necessitate admission to hospital. Furthermore, this composite endpoint has been used in several previous studies in this area.^{12,13,29} Using our composite endpoint, if patients can be identified as unlikely to need transfusion or haemostatic intervention, or die, then most doctors would be comfortable discharging them from the emergency department for outpatient management.

A weakness of our study was that we excluded patients who were already inpatients when they developed upper gastrointestinal bleeding. These patients are already known to be at high risk of poor outcome. In addition, 31% of patients did not undergo endoscopy. In three centres, there was a guideline to not admit patients with Glasgow Blatchford scores of either 1 or

less or 2 or less and aged less than 70 years, and many did not attend for outpatient endoscopy. In addition, some clinicians decided endoscopy was not required after an apparent trivial bleed, or was deemed inappropriate for some ill patients with multiple comorbidities. Interestingly, the UK national audit on upper gastrointestinal bleeding reported that 26% of patients notified to gastroenterologists did not have inpatient endoscopy.¹¹ We do not believe this had a major effect on our study findings, as 95% of patients were followed up for 30 days after presentation. Importantly, the usefulness of risk scores that require endoscopy is limited because endoscopy (if undertaken) is often delayed for up to 24 hours or more, whereas clinicians generally want risk stratification early after presentation.

A fairly large number of patients had missing values for full Rockall score (largely as a result of not undergoing endoscopy) or AIMS65 score (largely due to lack of measurement of albumin). Although 30 day mortality was similar in patients with and without missing values for both scoring systems, those with missing values had lower values of the other scores and less often underwent intervention, which could lead to an underestimation of the ability of the AIMS65 score to identify low risk patients. However, in the centre with the lowest missing values for AIMS65 score (Odense at 1.9%), subgroup analysis confirmed our main finding that the Glasgow Blatchford score had a higher discriminative ability to predict need for intervention or death compared with other risk scores, including AIMS65.

All data on patients managed as outpatients were included. In these outpatients, apart from performance of endoscopy they were otherwise assessed, investigated, and managed in the same way as all others and were followed up at 30 days using patient records. None of those managed as outpatients died and none required transfusion, surgery, or interventional radiology; endoscopic treatment was undertaken in one (a 28 year old with a minor Mallory-Weiss tear.). Therefore we do not believe that these policies had a major impact on our conclusions.

The power calculation for mortality was based on AIMS65 and full Rockall scores. Using our baseline assumptions, the number of patients in our study with full data on AIMS65 and full Rockall scores still provided 80% power to predict mortality. For our primary endpoint of intervention or death, using previous AUROC data for AIMS65 score of 0.7 and for the Glasgow Blatchford score of 0.8 (worst case value) and power of 90%, we needed only approximately 500 patients.³⁰

Table 5 | 30 day mortality at optimal score thresholds of Glasgow Blatchford score, AIMS65, PNED, and full Rockall score

Scoring system	Cut-off	No (%) of high risk patients*	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
AIMS65	≥2	672 (26.9)	65.8	76.2	18.0	96.6
PNED	≥4	1065 (37.6)	77.3	65.3	14.1	97.5
Admission Rockall	≥4	1130 (38.1)	78.6	65.0	14.3	97.6
Full Rockall	≥5	815 (40.5)	74.0	61.7	11.1	97.3
Glasgow Blatchford	≥5	1812 (61.8)	88.7	40.2	9.9	97.9

PPV=positive predictive value; NPV=negative predictive value.
*Classified as high risk according to risk scoring system.

Therefore we believe we achieved sufficient power in this study, which is the largest prospective study on this topic to date.

Comparison with other studies

A recent single centre retrospective study from Australia of 424 patients with upper gastrointestinal bleeding compared all these scores apart from PNEB.³¹ The researchers reported that the AIMS65 score was best at predicting mortality with a similar AUROC to ours, although lower than other smaller studies.^{16,19,29} They also found that AIMS65, Glasgow Blatchford, and full Rockall scores were similar at predicting their composite endpoint, which unlike ours did not include blood transfusion. However, we believe this is an important endpoint to assess, particularly for identification of low risk patients for possible outpatient management. A variety of low and high risk score thresholds have been suggested for clinical use by these and other small studies.

Published guidelines suggest outpatient management for patients with a Glasgow Blatchford score of 0.^{20,21,23} Some authors, however, have suggested that this low risk threshold could be extended to 1 or less or to 2 or less.^{13,14,22} We identified a Glasgow Blatchford score of 1 or less as the optimum score threshold to identify patients who would not require intervention or die, with a sensitivity of 98.6%. The accuracy of this threshold was consistently high across international sites, which comprised 19% of all patients in our study, compared with the 8.6% patients who had a Glasgow Blatchford score of 0. Of the patients with a Glasgow Blatchford score of 1 or less, all cause mortality was only 0.4%. This compares favourably with accepted low risk thresholds of commonly used scores for other medical conditions, including CURB65 for chest infection and the pulmonary embolism severity index.^{32,33} Other studies have reported more accurate prediction of death, rebleeding, or length of stay using these scores, but our study has the advantages of a multicentre design and large numbers of participants and suggests these outcomes are not well predicted by the studied scores.

The median length of hospital admission for a patient presenting with upper gastrointestinal bleeding in the UK is four or five days,^{3,12} and a recent study showed the mean in-hospital cost to be £2458 (\$3087; €2909) per patient.³⁴ Use of an accurate score within emergency departments or acute assessment units, to identify very low risk patients with upper gastrointestinal bleeding not requiring admission has clear benefits. Interestingly, sites varied in the proportion of very low risk patients defined by a Glasgow Blatchford score of 1 or less. The differences are probably due to different models of healthcare provision, including extent of primary care services. At the other end of the spectrum of severity of upper gastrointestinal bleeding, early identification of high risk patients who might benefit from urgent endoscopy or higher level care could help guide management. An observational study suggested that patients with a Glasgow Blatchford score greater than 12 have decreased mortality if endoscopy is undertaken less than 13 hours after presentation.¹⁹ Our data suggest

that a Glasgow Blatchford score of 7 or more and AIMS65 score of 2 or more have the highest combination of sensitivity and specificity for predicting endoscopic treatment or 30 day mortality, respectively, but the positive predictive value for both is low. Therefore the clinical utility of these scores to direct care in high risk patients seems limited.

Conclusions and policy implications

The Glasgow Blatchford score is accurate at predicting need for clinical intervention, or death in patients with upper gastrointestinal bleeding in all countries studied. A score of 1 or less is the optimum threshold for identifying very low risk patients suitable for outpatient management. A Glasgow Blatchford score of 7 or more is best at predicting need for endoscopic treatment, and a PNEB score of 4 or more and AIMS65 score of 2 or more are best at predicting mortality, although accuracy in predicting these endpoints is relatively low. No score seems accurate at predicting rebleeding or length of hospital stay. This information can help direct management of very low risk patients with upper gastrointestinal bleeding, but further studies using these, or new scores, are required to clarify their role in directing management of higher risk patients.

International Gastrointestinal Bleeding Consortium: P Dunne and J Fleming (Glasgow Royal Infirmary, Glasgow, UK); A Holloway, O Froud, J Sutton, N Michell, and E Pinchin (Royal Cornwall Hospital, Cornwall, UK); H S Ong, Y K Chin, and D S Koay (Singapore General Hospital, Singapore); T Steiner (Southern District Health Board, Dunedin Hospital, Dunedin, New Zealand).

Contributors: RA, LZ, ST, KW, CJK, SBL, PN, JF, AH, AF, JS, NM, EP, HSO, YKC, DSK and TS collected data. AJS and SBL wrote the paper with considerable input from LL, HRD, JHN, MS and IAM. All coauthors approved the final manuscript. AJS is the study guarantor.

Funding: This study was sponsored by NHS Greater Glasgow and Clyde Health Board, Scotland, UK. Data collection in Glasgow, Singapore, and Dunedin was funded by a health board endowment, hospital research, and guthealthnetwork grant, respectively. The funders had no input to the study design, data collection, or interpretation, writing of report, or submission for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that: (1) no authors have support from any company for the submitted work; (2) no authors have relationships with any company that might have an interest in the submitted work (3) no author, their spouses, partners, or children have financial relationships that may be relevant to the submitted work and (4) no authors have any non-financial interests that may be relevant to the submitted work.

Ethical approval: This study was approved by the west of Scotland ethics committee (reference 14/WS/0012; project No 145837) and each centre obtained approval from their local research committee or review board. Owing to the observational, non-interventional nature of the study, the ethical committee (and local committees and boards) agreed that individual patient consent was not required.

Data sharing: Patient level data and/or full dataset available from the corresponding author AJS. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Transparency: The study guarantor (AJS) affirms that the manuscript is an honest, accurate and transparent account of the study being reported. No important aspects of the study have been omitted and no discrepancies identified.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

- 1 Rockall TA, Logan RFA, Devlin HB, Northfield TC. Incidence of and mortality from acute upper gastrointestinal haemorrhage in the UK. *BMJ* 1995;311:222-6. doi:10.1136/bmj.311.6999.222.
- 2 Blatchford O, Davidson LA, Murray WR, Blatchford M, Pell J. Acute upper gastrointestinal haemorrhage in west of Scotland: case ascertainment study. *BMJ* 1997;315:510-4. doi:10.1136/bmj.315.7107.510.
- 3 Hearnshaw SA, Logan RFA, Lowe D, Travis SP, Murphy MF, Palmer KR. Acute upper gastrointestinal bleeding in the UK: patient characteristics, diagnoses and outcomes in the 2007 UK audit. *Gut* 2011;60:1327-35. doi:10.1136/gut.2010.228437.
- 4 Rockall TA, Logan RFA, Devlin HB, Northfield TC. Risk assessment after acute upper gastrointestinal haemorrhage. *Gut* 1996;38:316-21. doi:10.1136/gut.38.3.316.
- 5 Marmo R, Koch M, Cipolletta L, et al. Italian registry on upper gastrointestinal bleeding (Progetto Nazionale Emorragie Digestive--PNED 2). 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;105:1284-91. doi:10.1038/ajg.2009.687.
- 6 Saeed ZA, Winchester CB, Carolyn B, Michaeletz PA, Woods KL, Graham DY. A scoring system to predict rebleeding after endoscopic therapy of non-variceal upper gastrointestinal hemorrhage. *Am J Gastroenterol* 1993;88:1842-9.
- 7 Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper gastrointestinal haemorrhage. *Lancet* 2000;356:1318-21. doi:10.1016/S0140-6736(00)02816-6.
- 8 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;74:1215-24. doi:10.1016/j.gie.2011.06.024.
- 9 Das A, Ben-Menachem T, Farooq FT, et al. Artificial neural network as a predictive instrument in patients with acute nonvariceal upper gastrointestinal hemorrhage. *Gastroenterology* 2008;134:65-74. doi:10.1053/j.gastro.2007.10.037.
- 10 Laursen SB, Hansen JM, Schaffalitzky de Muckadell OB. The Glasgow Blatchford score is the most accurate assessment of patients with upper gastrointestinal hemorrhage. *Clin Gastroenterol Hepatol* 2012;10:1130-1135.e1. doi:10.1016/j.cgh.2012.06.022.
- 11 Hearnshaw SA, Logan RFA, Lowe D, Travis SPL, Murphy MF, Palmer KR. Use of endoscopy for management of acute upper gastrointestinal bleeding in the UK: results of a nationwide audit. *Gut* 2010;59:1022-9. doi:10.1136/gut.2008.174599.
- 12 Stanley AJ, Ashley D, Dalton HR, et al. Outpatient management of patients with low-risk upper-gastrointestinal haemorrhage: multicentre validation and prospective evaluation. *Lancet* 2009;373:42-7. doi:10.1016/S0140-6736(08)61769-9.
- 13 Laursen SB, Dalton HR, Murray IA, et al. Upper Gastrointestinal Hemorrhage International Consortium Upper Gastrointestinal Hemorrhage International Consortium. Performance of new thresholds of the Glasgow Blatchford score in managing patients with upper gastrointestinal bleeding. *Clin Gastroenterol Hepatol* 2015;13:115-21.e2. doi:10.1016/j.cgh.2014.07.023.
- 14 McLaughlin C, Vine L, Chapman L, et al. The management of low-risk primary upper gastrointestinal haemorrhage in the community: a 5-year observational study. *Eur J Gastroenterol Hepatol* 2012;24:288-93. doi:10.1097/MEG.0b013e32834febef.
- 15 Lim LG, Ho KY, Chan YH, et al. Urgent endoscopy is associated with lower mortality in high-risk but not low-risk nonvariceal upper gastrointestinal bleeding. *Endoscopy* 2011;43:300-6. doi:10.1055/s-0030-1256110.
- 16 Hyett BH, Abougergi MS, Charpentier JP, et al. The AIMS65 score compared with the Glasgow-Blatchford score in predicting outcomes in upper GI bleeding. *Gastrointest Endosc* 2013;77:551-7. doi:10.1016/j.gie.2012.11.022.
- 17 Stanley AJ, Dalton HR, Blatchford O, et al. Multicentre comparison of the Glasgow Blatchford and Rockall Scores in the prediction of clinical end-points after upper gastrointestinal haemorrhage. *Aliment Pharmacol Ther* 2011;34:470-5. doi:10.1111/j.1365-2036.2011.04747.x.
- 18 Pang SH, Ching JY, Lau JYW, Sung JY, Graham DY, Chan FKL. Comparing the Blatchford and pre-endoscopic Rockall score in predicting the need for endoscopic therapy in patients with upper GI hemorrhage. *Gastrointest Endosc* 2010;71:1134-40. doi:10.1016/j.gie.2010.01.028.
- 19 Yaka E, Yilmaz S, Doğan NÖ, Pekdemir M. Comparison of the Glasgow-Blatchford and AIMS65 scoring systems for risk stratification in upper gastrointestinal bleeding in the emergency department. *Acad Emerg Med* 2015;22:22-30. doi:10.1111/acem.12554.
- 20 Barkun AN, Bardou M, Kuipers EJ, et al. International Consensus Upper Gastrointestinal Bleeding Conference Group. International consensus recommendations on the management of patients with nonvariceal upper gastrointestinal bleeding. *Ann Intern Med* 2010;152:101-13. doi:10.7326/0003-4819-152-2-201001190-00009.
- 21 National Institute for Health and Clinical Excellence. *Acute Upper Gastrointestinal Bleeding: management. Clinical Guideline 141*. NICE, 2012.
- 22 Gralnek IM, Dumonceau JM, Kuipers EJ, et al. Diagnosis and management of nonvariceal upper gastrointestinal hemorrhage: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2015;47:a1-46. doi:10.1055/s-0034-1393172.
- 23 Laine L, Jensen DM. Management of patients with ulcer bleeding. *Am J Gastroenterol* 2012;107:345-60, quiz 361. doi:10.1038/ajg.2011.480.
- 24 de Franchis R, Baveno V Faculty. Revising consensus in portal hypertension: report of the Baveno V consensus workshop on methodology of diagnosis and therapy in portal hypertension. *J Hepatol* 2010;53:762-8. doi:10.1016/j.jhep.2010.06.004.
- 25 Villanueva C, Colomo A, Bosch A, et al. Transfusion strategies for acute upper gastrointestinal bleeding. *N Engl J Med* 2013;368:11-21. doi:10.1056/NEJMoa1211801.
- 26 Laine L, Spiegel B, Rostom A, et al. Methodology for randomized trials of patients with nonvariceal upper gastrointestinal bleeding: recommendations from an international consensus conference. *Am J Gastroenterol* 2010;105:540-50. doi:10.1038/ajg.2009.702.
- 27 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45. doi:10.2307/2531595.
- 28 Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.
- 29 Abougergi MS, Charpentier JP, Bethea E, et al. A prospective multicentre study of the AIMS65 score compared with the Glasgow Blatchford score in predicting upper gastrointestinal hemorrhage outcomes. *J Clin Gastroenterol* 2016;50:464-9. doi:10.1097/MCG.0000000000000395.
- 30 Palmer AJ, Moroni F, McLeish S, et al. P0440 A comparison of the Glasgow Blatchford and AIMS65 scores in predicting need for clinical intervention and mortality in acute non-variceal upper GI bleeds: a retrospective cohort study. *United European Gastroenterol J* 2014;2(Suppl):A132-605. doi:10.1177/2050640614548980.
- 31 Robertson M, Majumdar A, Boyapati 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;83:1151-60. doi:10.1016/j.gie.2015.10.021.
- 32 Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58:377-82. doi:10.1136/thorax.58.5.377.
- 33 Aujesky D, Roy PM, Verschuren F, et al. Outpatient versus inpatient treatment for patients with acute pulmonary embolism: an international, open-label, randomised, non-inferiority trial. *Lancet* 2011;378:41-8. doi:10.1016/S0140-6736(11)60824-6.
- 34 Campbell HE, Stokes EA, Bargo D, et al. Costs and quality of life associated with acute upper gastrointestinal bleeding in the UK: cohort analysis of patients in a cluster randomised trial. *BMJ Open* 2015;5:e007230doi:10.1136/bmjopen-2014-007230

Supplementary information: Supplementary tables 1 to 5.