

OPEN

A Novel Easy-to-Use Prediction Scheme for Upper Gastrointestinal Bleeding

Cologne-WATCH (C-WATCH) Risk Score

Vera Hoffmann, MD, Henrik Neubauer, Julia Heinzler, Anna Smarczyk, Martin Hellmich, DSc, Andrea Bowe, Fabian Kuetting, MD, Muenevver Demir, MD, Agnes Pelc, MD, Sigrid Schulte, MD, Ullrich Toex, MD, Dirk Nierhoff, MD, and Hans-Michael Steffen, MD

Abstract: Acute upper gastrointestinal bleeding (UGIB) is the leading indication for emergency endoscopy. Scoring schemes have been developed for immediate risk stratification. However, most of these scores include endoscopic findings and are based on data from patients with nonvariceal bleeding. The aim of our study was to design a pre-endoscopic score for acute UGIB—including variceal bleeding—in order to identify high-risk patients requiring urgent clinical management.

The scoring system was developed using a data set consisting of 586 patients with acute UGIB. These patients were identified from the emergency department as well as all inpatient services at the University Hospital of Cologne within a 2-year period (01/2007–12/2008). Further data from a cohort of 322 patients who presented to our endoscopy unit with acute UGIB in 2009 served for external/temporal validation.

Clinical, laboratory, and endoscopic parameters, as well as further data on medical history and medication were retrospectively collected from the electronic clinical documentation system.

A multivariable logistic regression was fitted to the development set to obtain a risk score using recurrent bleeding, need for intervention (angiography, surgery), or death within 30 days as a composite endpoint. Finally, the obtained risk score was evaluated on the validation set.

Only C-reactive protein, white blood cells, alanine-aminotransferase, thrombocytes, creatinine, and hemoglobin were identified as significant predictors for the composite endpoint. Based on the regression coefficients of these variables, an easy-to-use point scoring scheme (C-WATCH) was derived to estimate the risk of complications from 3% to 86% with an area under the curve (AUC) of 0.723 in the development set and 0.704 in the validation set. In the validation set, no patient in the identified low-risk group (0–1 points), but 38.7% of patients in the high-risk group (≥ 2 points) reached the composite endpoint.

Our easy-to-use scoring scheme is able to distinguish high-risk patients requiring urgent endoscopy, from low-risk cases who are suitable candidates for outpatient management or in whom endoscopy may be postponed. Based on our findings, a prospective validation of the C-WATCH score in different patient populations outside the university hospital setting seems warranted.

(*Medicine* 94(38):e1614)

Abbreviations: μl = microlitre, AIMS65 = Albumin $< 3\text{mmg/dL}$, international normalized ratio > 1.5 , systolic blood pressure $< 90\text{ mmHg}$, altered mental status age > 65 years, ALT = alanine aminotransferase, AST = aspartate aminotransferase, AUC = area under the curve, CI = confidence interval, CRP = C-reactive protein, C-WATCH = creatinine, white blood cells, alanine-aminotransferase, thrombocytes, C-reactive protein, hemoglobin, df = degrees of freedom, dL = decilitre, eg = example given, g = gram, GBS = Glasgow–Blatchford score, HL = Hosmer–Lemeshow, L = litre, LLN = lower limits of normal, MAD = maximum absolute difference, MCMC = Markov chain Monte Carlo, mg = milligram, mmHg = millimeter of mercury, n = number, OR = Odds ratio, p = p-value, PPIs = proton pump inhibitors, ref = reference value, ROC = receiver operating characteristic, SSRIs = selective serotonin reuptake inhibitors, U - units = units, UGIB = upper gastrointestinal bleeding, UK = United Kingdom, ULN = upper limits of normal.

INTRODUCTION

Acute upper gastrointestinal bleeding (UGIB) is a life-threatening condition with varying reports of incidence up to 47.7/100,000 persons annually, for example, in 2000 in the Netherlands.¹ The most frequent causes are peptic ulcers (46%), followed by esophagitis and erosions. Variceal bleeding has been found in 9% in the study of van Leerdam,¹ and in 11% in a United Kingdom (UK) study;² however, it is the major reason for bleeding in cirrhotic patients (50–60%).³ The risk of rebleeding is especially high in those suffering from peptic ulcers and variceal bleeding.³ Overall, mortality has decreased during the last decades from 14% in 1995 to 10% in 2007 in the United Kingdom.^{2,4}

Acute UGIB is an emergency that may need early endoscopic treatment; therefore, immediate triage of the patients is required.⁵ According to international consensus recommendations on the management of patients with nonvariceal UGIB, validated prognostic scores should be used in order to differentiate between high-risk patients in need of urgent endoscopic intervention and low-risk patients, who may be suitable candidates for outpatient management.⁶ Whereas there is evidence, that some of these low-risk patients could safely be managed in an

Editor: Mohamad Imam.

Received: May 20, 2015; revised: July 22, 2015; accepted: August 22, 2015.

From the Clinic for Gastroenterology and Hepatology, University Hospital of Cologne, Kerpener Str. 62, Köln, Germany (HV, NH, HJ, SA, BA, KF, DM, PA, SS, TU, ND, SHM); Institute of Medical Statistics, Informatics and Epidemiology, University of Cologne, Kerpener Str. 62, Köln, Germany (HM).

Correspondence: Vera Hoffmann, Clinic for Gastroenterology and Hepatology, University Hospital of Cologne, Kerpener Str. 62, Köln, Germany (e-mail: Vera.Hoffmann_@uk-koeln.de).

VH, DN, and HMS equally contributed to this study.

Supplemental Digital Content is available for this article.

The authors have no funding and conflicts of interest to disclose.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0, where it is permissible to download, share and reproduce the work in any medium, provided it is properly cited. The work cannot be changed in any way or used commercially.

ISSN: 0025-7974

DOI: 10.1097/MD.0000000000001614

outpatient setting,^{5,7–10} many patients with low risk for further complications are hospitalized, nonetheless. With risk-adapted selection of patients who are unlikely to get harmed due to early discharge, costs and resources could be conserved.

Several risk scoring schemes or systems evaluating acute UGIB have already been published. However, many of these predictive schemes rely on endoscopic findings and are, therefore, unsuitable for early classification of patients shortly after presentation.^{11–20} In addition, the development of most risk scores is based solely on data from patients with nonvariceal bleeding and thus cannot be used to evaluate a mixed patient population.^{13–19} Even with the well-established Glasgow–Blatchford score (GBS), Lahiff et al were unable to predict the outcome in a study with a mixed patient population of variceal and nonvariceal bleeding.²¹ In the emergency setting of UGIB it is generally difficult to differentiate immediately between nonvariceal and variceal bleeding. At the same time, mortality in variceal bleeding is high (around 15%);² thus a reliable risk stratification of all patients with UGIB would be very helpful.

According to a recently published statement by Stanley et al, a high-quality score should achieve the following: easy to calculate, accurate with respect to relevant outcomes, and capable of early, ideally before endoscopy, risk assessment.²² Therefore, we aimed at developing a point scoring scheme for all patients with acute UGIB, including those with variceal bleeding that predicts the clinically meaningful composite endpoint of rebleeding, need for nonendoscopic intervention (angiography and surgery), or death within 30 days. The scoring scheme should consist of easily obtainable laboratory values in order to reliably differentiate between high-risk patients who require hospitalization or immediate upper endoscopy from low-risk patients that are suitable candidates for outpatient management or in whom endoscopy may safely be postponed.

PATIENTS AND METHODS

Data from all patients who had undergone upper endoscopy in the Clinic of Gastroenterology and Hepatology at the University Hospital of Cologne with clinically suspected acute UGIB (eg, hematemesis, melena, hematochezia, or fall in hemoglobin) within a 2-year period (1/1/2007–12/31/2008) were analyzed retrospectively and served as a development set. There were 363 patients from all in-hospital services and 258 emergency room patients with a mean age of 64.3 ± 15.1 years. A population of 165 patients from all in-hospital and 175 emergency room patients who fulfilled the same criteria as the development set and were seen between 01/01/2009 and 12/31/2009 served as an external/temporal validation set. Patients who primarily presented to another hospital were excluded from our study. Thus, a total of 586 and 286 patients were included in the development and validation set, respectively.

Ethical approval was not required for this study since the study has a retrospective epidemiological design and anonymization of personal data was realized.

The following data were extracted from the clinical documentation system (Orbis, Agfa HealthCare, Bonn, Germany): age, gender, past medical history, medication, laboratory values, endoscopic findings, need for intensive care, length of hospital stay, and clinical course. Relevant past medical history included former ulcers or UGIB, pulmonary disease, arterial hypertension, diabetes, stroke, cardiac disease, dementia, renal failure, hepatitis, liver failure, arthritis, previous surgery, status post

reanimation, sepsis, trauma, current or former malignant disease, smoking-, and alcohol abuse. Medications with known bleeding risks, such as antiplatelet drugs, nonsteroidal anti-inflammatory drugs, selective serotonin reuptake inhibitors (SSRIs), steroids, heparin, and vitamin K antagonists were also documented, as well as gastric acid inhibition with proton pump inhibitors (PPIs) or H₂-blockers. Documented clinical parameters included symptoms associated with UGIB such as hematemesis, coffee-ground vomiting, tarry stool, hematochezia, syncope, and paleness, as well as signs of hypovolaemic shock and nasogastric tube placement with or without bloody aspirate. In order to avoid effects of blood transfusions the documented laboratory values were selected from the first blood sample taken after presentation with clinically suspected UGIB. The cutoffs for the laboratory values are the upper (ULN) or lower (LLN) limits of normal, respectively, as given by the Institute for Clinical Chemistry at the University Hospital of Cologne. The relevant cutoffs were as follows: CRP (5 mg/dL), white blood cell count (11,300/ μ L), alanine-aminotransferase (ALT, 35 U/L for women, 50 U/L for men), thrombocytes (150,000/ μ L), creatinine (0.9 mg/dL for women, 1.1 mg/dL for men), and hemoglobin (12 g/dL for women, 14 g/dL for men). For hemoglobin and thrombocytes a second cutoff was defined. For hemoglobin 10 mg/dL was chosen since often clinical symptoms appear below this value. The cutoff for thrombocytes of 50,000/ μ L was chosen since this is the typical limit used for endoscopic interventions. Table 1 gives an overview of relevant characteristics of the study population.

Emergency endoscopy was performed on a 24 h/7 days a week schedule by a team of 6 experienced physicians. Lesions were rated by using the Forrest classification.²³ Spurting or oozing fresh blood was considered as active bleeding, whereas clots or old blood were documented as signs of bleeding. In addition, all macroscopic findings were documented, that is, esophagitis, esophageal, or gastric varices, Mallory–Weiss lesion, gastritis, erosions, angiodysplasia, Dieulafoy-lesion, and ulcers with localization, suspicion of malignancy, and fresh or old blood without source of bleeding. Therapeutic endoscopic interventions such as injections with suparenine, ethoxysclerol, fibrin- or histoacryl glue, hemoclips, laser or argonplasma-coagulation were left to the discretion of the treating physician, as well as infusion of plasma expander, PPIs, somatostatin or number of packed red cell transfusions that were recorded.

The need for angiographic or surgical interventions was also extracted from the digital patient files for the 30 days after the index hemorrhage. Rebleeding was defined as the occurrence of fresh or old blood seen during a second endoscopy that had been ordered also within 30 days on clinical suspicion for rebleeding, for example, decrease in hemoglobin > 2 g/dL within 24 h by the treating physician. Vital status within 30 days after the index hemorrhage was confirmed by contacting the appropriate civil registry offices.

Statistical Analysis

In the first step, based on the development set the association of the composite outcome and each candidate variable was assessed one by one via cross-tabulation (Pearson's chi-square test, odds ratio). Continuous variables were dichotomized according to conventional clinical threshold values. In the second step, the candidate variables showing a statistically significant association (ie $P \leq 0.05$) were entered in a multivariable logistic regression equation with backward stepwise

TABLE 1. Characteristics of the Study Population

Characteristics	Development Set (n = 586)			Validation Set (n = 286)			P-Value*
	Count	Mean	Standard Deviation	Count	Mean	Standard Deviation	
Age (years)	586	64.3	15.1	286	64.5	14.2	0.769
Sex							
	Female	203 (34.6%)		93 (32.5%)			0.534
	Male	383 (65.4%)		193 (67.5%)			
Rebleeding							0.591
	Yes	96 (16.4%)		51 (17.8%)			
	No	490 (83.6%)		235 (82.2%)			
Angiography							0.179
	Yes	9 (1.5%)		1 (0.3%)			
	No	577 (98.5%)		285 (99.7%)			
Surgery							0.117
	Yes	11 (1.9%)		1 (0.3%)			
	No	575 (98.1%)		285 (99.7%)			
Death (n = 560)							0.913
	Yes	118 (21.1%)		56 (20.7%)			
	No	442 (78.9%)		214 (79.3%)			
C-reactive protein (mg/dl) (ref: < 5.0 mg/l)	526	66.3	78.3	214	73.7	83.0	0.046
White blood cells (count x1 ⁹ /l) (ref: 4.4–11.3 x1 ⁹ /l)	586	10.9	6.6	286	11.9	11.8	0.745
ALT (U/l) (ref: 35 U/L women, > 50 U/L men)	487	69.4	224.3	222	69.3	402.8	0.482
Thrombocytes (count x1 ⁹ /l) (ref: 150–400)	586	216.3	132	286	196.3	133.8	0.013
Creatinine (mg/dl) (ref: > 0.9 women, > 1.1 men)	546	2.2	3.3	246	2.2	2.1	0.906
Hemoglobin (g/dl) (ref: < 12 women, < 14 men)	586	9.7	2.4	286	9.3	2.3	0.001

ALT = alanine aminotransferase, ref = reference value.

elimination (drop if $P > 0.05$). Beforehand, multiple imputation using an iterative Markov chain Monte Carlo (MCMC) method (fully conditional specification, 5 imputation sets) was applied to candidate variables with at least 80% valid observations. In sensitivity analyses, (1) impact of a greater percentage of missing values (eg 30%, 50%), and (2) pairwise interaction of covariates were explored. Alternative risk score models were either rejected due to lower discriminatory capacity or interaction terms not reaching statistical significance. The pooled regression coefficients were used to calculate predicted values (ie complication risk). In the third step, the model coefficients were scaled to integer numbers (division by the smallest coefficient, adjustment, rounding) to get a simple point scoring scheme.²⁴ In the fourth step, the final score was evaluated on a separate validation set, which had not been touched during development. Discrimination of the score was assessed by the receiver operating characteristic curve and the area under the curve, calibration (goodness of fit) of the score was measured by the Hosmer–Lemeshow statistic.²⁵ Moreover, the corresponding observed and expected proportions of events were plotted by “decile of risk.” Since subjects with the same covariate pattern were kept in the same “decile,” only 6 groups were obtained for the point score. All calculations were carried out with the software SPSS Statistics 22 (IBM Corp., Armonk, NY).

RESULTS

The development set consisted of 203 women (34.6 %) and 383 men (65.4 %) (Table 1). The mean age was 64.3 ± 15.1 years. In this group, 77 patients (13.1 %) showed esophageal or

gastric varices during endoscopy, whereas 39 patients of these patients showed variceal bleeding (6.7 %). The 3 most frequent lesions identified were erosions (39.8 %), ulcers (32.8 %), and esophagitis (22.4 %). Endoscopic findings are listed in detail in Table 2. Active bleeding was observed in 141 patients (24.1 %), recurrent bleeding occurred in 96 patients (16.4 %), and a total of 118 patients died within the next 30 days (21.1 %).

The following variables were not considered due to a large proportion of missing values (ie > 20%): albumin, alkaline phosphatase, bilirubin, cholinesterase, aspartate aminotransferase (AST), blood pressure, heart rate, cutaneous signs of liver disease, liver failure, septicaemia, pneumonia, stroke, renal failure, need for dialysis, coronary heart disease, active malignant disease, concomitant medication with PPIs, history of former bleeding, ulcer, or malignant disease.

In bivariate analysis, 17 variables (age in years and 16 dichotomized variables) significantly predicted the composite endpoint (see Supplement, <http://links.lww.com/MD/A432>). Using multivariable logistic regression, we identified 6 out of these 17 parameters, that were predictive for the composite endpoint of recurrent bleeding, need for intervention such as angiography or emergency surgery and death within 30 days after the index hemorrhage: CRP > 5 mg/dL (odds ratio [OR] 2.09, 95% confidence interval [CI] 1.06–4.15), white blood cell count > 11,300 per μ L, (OR 2.46, 95% CI 1.66–3.66), ALT > 35 U/L for women or > 50 U/L for men (OR 1.85, 95% CI 1.04–3.29), thrombocytes \geq 50,000 and < 150,000 per μ L (OR 1.84, 95% CI 1.2–2.82; and OR 2.82, 95% CI 1.36–5.85, respectively), creatinine > 0.9 mg/dL for women or > 1.1 mg/dL for men (OR 2.15, 95% CI 1.28–3.62), and hemoglobin < 10 g/

TABLE 2. Endoscopic Findings

Endoscopic Findings		Count	%
Signs of bleeding*	Yes	178	30.4
	No	408	69.6
Active bleeding	Yes	141	24.1
	No	445	75.9
Varices	Yes	77	13.1
	No	509	86.9
Variceal bleeding	Yes	39	6.7
	No	547	93.4
Esophagitis	Yes	131	22.4
	No	455	77.7
Mallory–Weiss lesions	Yes	14	2.4
	No	572	97.6
Gastritis	Yes	89	15.2
	No	497	84.8
Erosions	Yes	233	39.8
	No	353	60.2
Ulcer	Yes	192	32.8
	No	394	67.2
Suspicion of malignancy	Yes	19	32.8
	No	394	67.2

* Fresh or old blood during endoscopy.

dL and < 12 g/dL for women, < 10 g/dL and < 14 g/dL for men (OR 2.98, 95% CI 1.21–7.34; < 10 g/dL, OR 2.54, 95% CI 1–6.43).

From the multivariable logistic regression equation “C-WATCH MODEL”

Complication risk = $1 / \{1 + \exp[-0.766 * (\text{insert 1 if creatinine is } > \text{ULN, else 0}) - 0.739 * (\text{insert 1 if CRP is } > \text{ULN, else 0}) - 0.900 * (\text{insert 1 if white blood cell count } > \text{ULN, else 0}) - 0.615 * (\text{insert 1 if ALT } > \text{ULN, else 0}) - 1.090 * (\text{insert 1 if hemoglobin } < 10 \text{ g/dL, else 0}) - 0.930 * (\text{insert 1 if hemoglobin } \geq 10 \text{ g/dL and } < 12 \text{ g/dL for women } / \geq 10 \text{ g/dL and } < 14 \text{ g/dL for men, else 0}) - 1.038 * (\text{insert 1 if thrombocytes } < 50,000, \text{ else 0}) - 0.609 * (\text{insert 1 if thrombocytes } \geq 50,000 \text{ and } < 150,000 \text{ per } \mu\text{L, else 0}) + 3.564\}$.

We derived a simple point scoring scheme “Cologne-WATCH score” (C-WATCH score) by adequate rounding of the regression coefficients (Figure 1).

The discriminatory power of “model” and “score” is high in the development set with an area under the ROC curve (AUC) of 0.739 (95% CI 0.692–0.786) and 0.723 (95% CI 0.676–0.770), respectively. Calibration of the model was good with a Hosmer–Lemeshow (HL) statistic of 6.47 (8 degrees of freedom [df], $p = 0.595$) and a maximum absolute difference (MAD) of 10% in observed and expected events over “deciles of risk.” Since the score is a simplification of the model, we expected a slightly worse calibration (HL statistic of 13.06 (4 df, $P = 0.011$); however the MAD was still 10%. In the validation set, the model achieved an AUC of 0.700 (95% CI 0.638–0.761) and the score of 0.704 (95% CI 0.642–0.765). Calibration was somewhat worse for the model (HL statistic of 22.13, 8 df, $P = 0.005$, MAD of 23%), however good for the score (HL statistic of 9.08, 4 df, $P = 0.059$, MAD 10%). The ROC curves and calibration plots are shown in Figure 2.

A total of 34% of patients in the development set and 36% of patients in the validation set experienced a complication. In the development set, only 1 out of 6 patients rated with zero

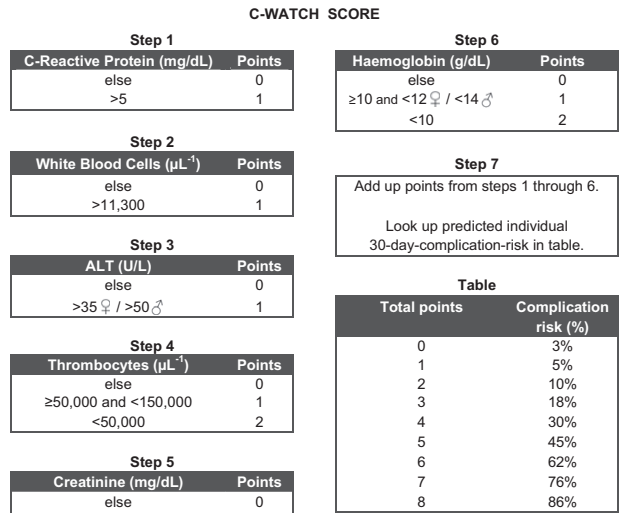


FIGURE 1. C-WATCH score scheme of the 7 steps for calculating the C-WATCH score.

points and none out of 38 patients rated with 1 point suffered a complication during follow up (one patient experienced a rebleeding event). In the validation set, 10 patients were rated with a score of zero and 9 patients with a score of 1. No complication occurred in both categories. Thus, 8% in the development set and 7% in the validation set had a score of 0 to 1 and can be defined as low-risk patients with no relevant risk of complications (see Table 3). A higher number of the score is associated with an increasing risk for the event of any complication (3% to 86%, see Figure 1), for example, complications were observed in 12 of 14 high-risk patients with a score of 7. Nota bene, in the validation set all patients with variceal bleeding scored ≥ 2 points and, thus, would have been managed as high-risk patients.

DISCUSSION

Acute UGIB is the most important emergency situation in gastroenterology. Due to the fact that it is associated with a high mortality of ~10%, most of the patients are monitored and remain hospitalized for several days.² High-risk patients with hemodynamic impairment undoubtedly require immediate endoscopy whereas in many cases upper endoscopy within 24 h may suffice.⁵ However, there is a subgroup of patients with a low risk of active bleeding or complications, which do not need hospitalization or early endoscopic intervention.⁵ Identifying these low-risk patients would help to save resources and reduce cost by managing them as outpatients. Most published risk scores have been developed in study populations with nonvariceal bleeding and hence are not valid for the evaluation of patients with varices.^{13–19}

We designed a pre-endoscopic risk score for all patients with UGIB—including variceal bleeding—from a retrospective analysis of 586 patients from inpatient services as well as emergency room visits who presented to our clinic within a 2-year period. By logistic regression, we were able to identify 6 laboratory parameters which proved to be valid to predict recurrent bleeding, need for surgery or angiographic intervention, and death within 30 days: CRP, white blood cell count, alanine-aminotransferase, thrombocytes, creatinine, and hemoglobin: the C-WATCH score.

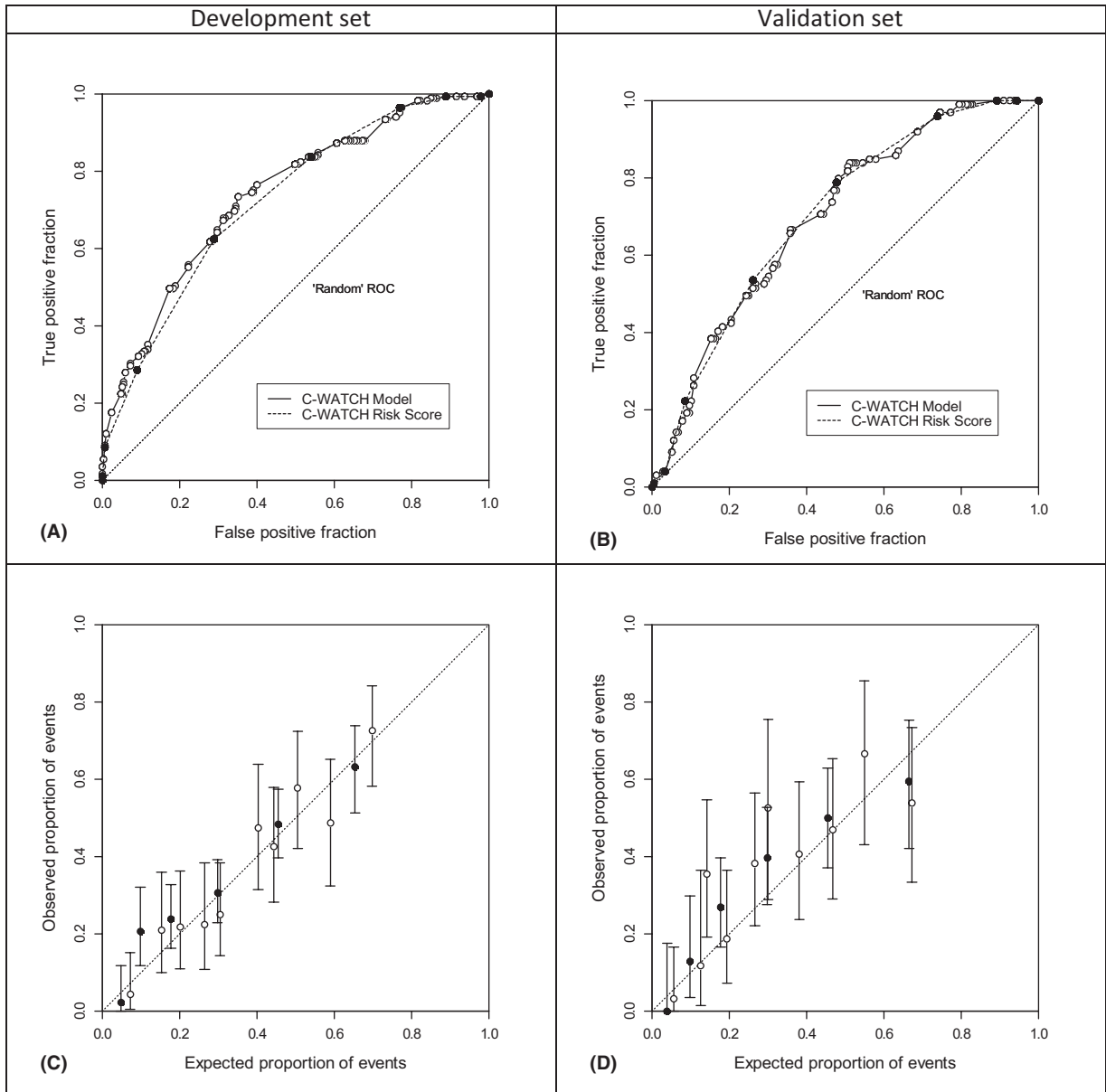


FIGURE 2. Area under ROC curve discrimination between patients with complication (yes/no) in the development set (A: C-WATCH model, area under ROC curve 0.739, and C-WATCH score, AUC 0.723) and the validation set (B: AUC 0.700 and 0.704, respectively). Calibration plots of both models (white points) and risk score (black points) in the development set (C) and validation set (D). Whiskers denote 95% confidence intervals (Clopper–Pearson). ROC = receiver operating characteristic

All components of C-WATCH score are meaningful from a clinical or pathophysiological point of view in connection with UGIB.^{26–28}

The C-WATCH score allows us to separate low-risk patients with <2 points (0% complications in the validation set), who can be managed on an outpatient basis from high-risk patients with ≥ 2 points (38.7% complications in the validation set) who should be monitored in hospital. When combining the development and validation sets, only one out of 63 patients rated as low-risk was misclassified when the C-WATCH score was applied. This patient had an episode of rebleeding within 30 days of follow up. The overall diagnostic accuracy of this scoring scheme was acceptable with an AUC of 0.723 and

0.704 for the development and validation set, respectively (see Figure 2 and 3). Although it seems reasonable to postpone endoscopy by 24–48 h in patients with a score of 0–1, it will depend on the desired safety margin and clinical judgment whether in the high-risk group endoscopy should be performed within 12–24 hours (score 2–3 corresponding to a 10–18% complication risk) or immediately (score ≥4 corresponding to a 30–86% complication risk).

Most scoring systems have been based on smaller study populations with 108 to 391 patients included.^{12,15–18,29–33} Also, in a recently published systematic review of prediction scores, the median number of patients in the included studies accounted to 248 patients.³⁴ We were able to include nearly

TABLE 3. Complications in Relation to the C-WATCH Score

C-WATCH Score	Development Set (n = 566*)		Validation Set (n = 275†)	
	No Complication	Complication	No Complication	Complication
0	6 (1.6%)	1 (0.5%)	10 (5.7%)	0 (0.0%)
1	38 (10.2%)	0 (0.0%)	9 (5.1)	0 (0.0%)
2	54 (14.5%)	14 (7.2%)	27 (15.3%)	4 (4.0%)
3	86 (23.1%)	27 (13.9%)	46 (26.1%)	17 (17.2%)
4	93 (25.0%)	41 (21.1%)	38 (21.6%)	25 (25.3%)
5	67 (18.0%)	63 (32.5%)	31 (17.6%)	31 (31.3%)
6	26 (7.0%)	34 (17.5%)	9 (5.1%)	18 (18.2%)
7	2 (0.5%)	12 (6.2%)	5 (2.8%)	3 (3.0%)
8	0 (0.0%)	2 (1.0%)	1 (0.6%)	1 (1.0%)
Total	372 (65.7%)	194 (34.3%)	176 (64.0%)	99 (36%)

* 26 missing death status, 6 of which had a rebleeding.

† 16 missing death status, 5 of which had a rebleeding.

twice as many patients (n = 586) in our development set, resulting in an improved reliability of the prediction model.

Another important advantage of our prediction tool is the inclusion of variceal hemorrhage. In an emergency situation, discrimination of this high-risk group from those presenting with a different source of bleeding can be challenging. However, the prevalence of variceal bleeding has been found to range from 9% to 11% which is comparable to the rate of 7% in our study population.^{1,2} Patients with variceal hemorrhage show a high mortality (38% in our development set). A clinically useful predictive tool should include these complication prone patients. In the validation set all patients with variceal bleeding scored ≥ 2 points and, thus, would have been managed as high-risk patients.

Many of the established scoring systems try to predict the need for urgent endoscopic intervention or focus on the risk assessment of a poor clinical outcome in high-risk patients.^{12–14,16–18,29,30,35} The predictive power of these diagnostic tools to identify low-risk patients has not been adequately studied. The meta-analysis by de Groot et al recommends four scoring systems to predict mortality, rebleeding, need for intervention or poor outcome. However, none has been suggested to identify low-risk patients.³⁴

Early endoscopy is able to identify low-risk patients as candidates for outpatient management and skipping urgent endoscopy in these patients is not associated with an increase in mortality.^{5,36} As the C-WATCH score can detect low-risk patients even without performing upper endoscopy, its clinical use as a prediction tool is highly attractive. Identifying patients as suitable candidates for early discharge would facilitate cost-effective outpatient management without jeopardizing patient safety. Analyses of the GBS and the pre-endoscopic Rockall-Score in this regard have reached contradictory results. Stanley et al reported a safe outpatient management following risk stratification with the GBS: Of 676 studied patients, 105 (16%) were classified with a score of zero and none of these were misclassified in regard to the endpoint of death or need for intervention.³⁷ In contrast, Meltzer et al reported an insufficient sensitivity: 63 patients from 690 were classified with a GBS of zero points; 15 (24%) of these were admitted to the hospital and 2 patients (13%) needed endoscopic treatment.³⁸ The pre-endoscopic Rockall score showed even worse results with regard to this question: 55% of the patients with a Rockall

score of zero were admitted to the hospital and 16% needed endoscopic hemostasis.

The analysis of prediction scores for gastrointestinal bleeding by de Groot et al deals with 16 published scoring schemes.³⁴ These systems were evaluated using a rating scale (1–29 points) with defined criteria that a useful predictive scoring scheme should accomplish in order to be of high quality. The mean overall quality rating of the reviewed prediction tools was 16.5 points (ranging 9–25 points).³⁴ Using this quality rating the C-WATCH score attains 21 points which compares favorably to the recommended scores from the analysis by de Groot et al: that is the Blatchford score (25 points) to evaluate the need for intervention, the score of Villanueva et al for poor outcome (19 points), the score of Guglielmi et al for rebleeding (21 points), and the score of Chui et al for mortality (21 points).^{14,15,18,35} Criticized shortcomings of the appraised scoring schemes were the lack of external validation, the reference to a single endpoint, and the absence of an impact analysis. We performed a temporal validation in a further patient population and could demonstrate a similar predictive power with an AUC of 0.723 and 0.704, respectively. This is remarkable, considering statistically significant differences between analysis sets, for example, regarding CRP, hemoglobin, and thrombocytes (see Table 1). Furthermore our scoring scheme fulfills all of the prespecified criteria for an ideal prediction scheme from a recently published update on risk scoring systems in UGIB²²: The C-WATCH score can be calculated early after patient presentation even by someone without endoscopic experience, is easy to use by simply adding the points from routine laboratory parameters, and is accurate for the prediction of relevant outcomes with a high discriminatory power.

The C-WATCH score was derived with reference to a composite endpoint of hard clinical parameters including recurrent bleeding, need for angiography or surgery and 30 day mortality. We excluded endoscopic findings or interventions as part of the composite endpoint because they are not completely operator independent. However, our study has several limitations. Due to its retrospective design, data had to be retrieved from stored patient files with missing or incomplete data in several cases. In order to derive a practical and reliable scoring scheme, we included only variables that could be retrieved for at least 80% of patients. Thus, several potentially relevant parameters, for example, blood pressure, were not included

in our main analysis; however, all were considered in sensitivity analyses. There was no standardized re-assessment of laboratory parameters, for example, hemoglobin after a defined time period, in order to include the dynamic of changes into the prediction model. Also, the C-WATCH score does not take trends in patient clinical stability into account, for example, ongoing bloody vomiting or hemodynamic deterioration. Hence, clinical judgment is still essential and cannot be replaced by this scoring scheme.

With regard to our subgroup of variceal bleeding death within 6 weeks has been proposed as clinical endpoint by the Baveno V consensus workshop on portal hypertension.³⁹ However, the evidence for this expert recommendation remains unclear. Also, variceal bleeding is a small subgroup of only 6.7% in our cohort. For our main population with other source of bleeding a 6 weeks follow-up seems too long for acute bleeding and would not be helpful. Thus we decided to use a 30 days follow up like generally used in case of acute nonvariceal bleeding.⁴⁰

Another shortcoming of our study is the missing external validation in a patient population different from a tertiary care facility, as well as the validation in a prospective study. Finally, the C-WATCH score has not yet undergone an impact analysis and the benefit for patients in terms of outcome and survival remains to be proven in a prospective study; for example, in a direct comparison with the GBS as it has been done with the AIMS65 score by Yaka et al⁴¹

In conclusion, the C-WATCH score allows for an operator-independent, easy-to-use risk assessment in out- and inpatients with acute UGIB including variceal hemorrhage. Low-risk patients suitable for outpatient management or delayed upper endoscopy can be identified early after presentation and urgent or immediate upper endoscopy can be reserved for patients with high risk, depending on clinical judgment and the desired safety margin.

ACKNOWLEDGMENTS

This study was not funded by any external organization. All authors are associated with the University hospital of Cologne.

REFERENCES

- van Leerdam ME, Vreeburg EM, Rauws EA, 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;98:1494–1499.
- Hearnshaw SA, Logan RF, Lowe D, et al. Acute upper gastrointestinal bleeding in the UK: patient characteristics, diagnoses and outcomes in the 2007 UK audit. *Gut.* 2011;60:1327–1335.
- van Leerdam ME. Epidemiology of acute upper gastrointestinal bleeding. *Best Pract Res Clin Gastroenterol.* 2008;22:209–224.
- Rockall TA, Logan RF, Devlin HB, et al. Incidence of and mortality from acute upper gastrointestinal haemorrhage in the United Kingdom. Steering Committee and members of the National Audit of Acute Upper Gastrointestinal Haemorrhage. *BMJ.* 1995;311:222–226.
- 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;43:300–306.
- Barkun AN, Bardou M, Kuipers EJ, et al. International consensus recommendations on the management of patients with nonvariceal upper gastrointestinal bleeding. *Ann Intern Med.* 2010;152:101–113.
- Lai KC, Hui WM, Wong BC, et al. A retrospective and prospective study on the safety of discharging selected patients with duodenal ulcer bleeding on the same day as endoscopy. *Gastrointest Endosc.* 1997;45:26–30.
- Longstreth GF, Feitelberg SP. Outpatient care of selected patients with acute non-variceal upper gastrointestinal haemorrhage. *Lancet.* 1995;345:108–111.
- Cipolletta L, Bianco MA, Rotondano G, et al. Outpatient management for low-risk nonvariceal upper GI bleeding: a randomized controlled trial. *Gastrointest Endosc.* 2002;55:1–5.
- Robins GG, Sarwar MS, Armstrong MJ, et al. Evaluation of the need for endoscopy to identify low-risk patients presenting with an acute upper gastrointestinal bleed suitable for early discharge. *Postgrad Med J.* 2007;83:768–772.
- Rockall TA, Logan RF, Devlin HB, et al. Risk assessment after acute upper gastrointestinal haemorrhage. *Gut.* 1996;38:316–321.
- Pimpl W, Boeckl O, Waclawiczek HW, et al. Estimation of the mortality rate of patients with severe gastroduodenal hemorrhage with the aid of a new scoring system. *Endoscopy.* 1987;19:101–106.
- Marmo R, Koch M, Cipolletta L, et al. Predicting mortality in non-variceal upper gastrointestinal bleeders: validation of the Italian PNEED score and prospective comparison with the Rockall score. *Am J Gastroenterol.* 2010;105:1284–1291.
- Chiu PW, Ng EK, Cheung FK, et al. Predicting mortality in patients with bleeding peptic ulcers after therapeutic endoscopy. *Clin Gastroenterol Hepatol.* 2009;7:311–316quiz 253..
- Guglielmi A, Ruzzenente A, Sandri M, et al. Risk assessment and prediction of rebleeding in bleeding gastroduodenal ulcer. *Endoscopy.* 2002;34:778–786.
- Travis AC, Wasan SK, Saltzman JR. Model to predict rebleeding following endoscopic therapy for non-variceal upper gastrointestinal hemorrhage. *J Gastroenterol Hepatol.* 2008;23:1505–1510.
- Park KG, Steele RJ, Mollison J, et al. Prediction of recurrent bleeding after endoscopic haemostasis in non-variceal upper gastrointestinal haemorrhage. *Br J Surg.* 1994;81:1465–1468.
- Villanueva C, Balanzo J, Espinos JC, et al. Prediction of therapeutic failure in patients with bleeding peptic ulcer treated with endoscopic injection. *Dig Dis Sci.* 1993;38:2062–2070.
- Almela P, Benages A, Peiro S, et al. A risk score system for identification of patients with upper-GI bleeding suitable for outpatient management. *Gastrointest Endosc.* 2004;59:772–781.
- Imperiale TF, Dominitz JA, Provenzale DT, et al. Predicting poor outcome from acute upper gastrointestinal hemorrhage. *Arch Intern Med.* 2007;167:1291–1296.
- Lahiff C, Shields W, Cretu I, et al. Upper gastrointestinal bleeding: predictors of risk in a mixed patient group including variceal and nonvariceal haemorrhage. *Eur J Gastroenterol Hepatol.* 2012;24:149–154.
- Stanley AJ. Update on risk scoring systems for patients with upper gastrointestinal haemorrhage. *World J Gastroenterol.* 2012;18:2739–2744.
- Forrest JA, Finlayson ND, Shearman DJ. Endoscopy in gastrointestinal bleeding. *Lancet.* 1974;2:394–397.
- Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: the Framingham study risk score functions. *Stat Med.* 2004;23:1631–1660.
- Hosmer DW LS. *Applied Logistic Regression.* 2nd ed. New York: Wiley; 2000, 35–226.
- Tomizawa M, Shinozaki F, Hasegawa R, et al. Reduced hemoglobin and increased C-reactive protein are associated with upper gastrointestinal bleeding. *World J Gastroenterol.* 2014;20:1311–1317.
- Lee HH, Park JM, Lee SW, et al. C-reactive protein as a prognostic indicator for rebleeding in patients with nonvariceal upper gastrointestinal bleeding. *Dig Liver Dis.* 2015;47:378–383.

28. Cerqueira RM, Andrade L, Correia MR, et al. Risk factors for in-hospital mortality in cirrhotic patients with oesophageal variceal bleeding. *Eur J Gastroenterol Hepatol*. 2012;24:551–557.
29. 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;25:1125–1132.
30. Provenzale D, Sandler RS, Wood DR, et al. Development of a scoring system to predict mortality from upper gastrointestinal bleeding. *Am J Med Sci*. 1987;294:26–32.
31. Bordley DR, Mushlin AI, Dolan JG, et al. Early clinical signs identify low-risk patients with acute upper gastrointestinal hemorrhage. *JAMA*. 1985;253:3282–3285.
32. Corley DA, Stefan AM, Wolf M, et al. Early indicators of prognosis in upper gastrointestinal hemorrhage. *Am J Gastroenterol*. 1998;93:336–340.
33. Strate LL, Orav EJ, Syngal S. Early predictors of severity in acute lower intestinal tract bleeding. *Arch Intern Med*. 2003;163:838–843.
34. de Groot NL, Bosman JH, Siersema PD, et al. Prediction scores in gastrointestinal bleeding: a systematic review and quantitative appraisal. *Endoscopy*. 2012;44:731–739.
35. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper-gastrointestinal haemorrhage. *Lancet*. 2000;356:1318–1321.
36. Spiegel BM, Vakil NB, Ofman JJ. Endoscopy for acute nonvariceal upper gastrointestinal tract hemorrhage: is sooner better? A systematic review. *Arch Intern Med*. 2001;161:1393–1404.
37. Stanley AJ, Ashley D, Dalton HR, et al. Outpatient management of patients with low-risk upper-gastrointestinal haemorrhage: multi-centre validation and prospective evaluation. *Lancet*. 2009;373:42–47.
38. Meltzer AC, Burnett S, Pinchbeck C, et al. Pre-endoscopic Rockall and Blatchford scores to identify which emergency department patients with suspected gastrointestinal bleed do not need endoscopic hemostasis. *J Emerg Med*. 2013;44:1083–1087.
39. de Franchis R, Baveno VF. Revising consensus in portal hypertension: report of the Baveno V consensus workshop on methodology of diagnosis and therapy in portal hypertension. *J Hepatol*. 2010;53:762–768.
40. de Groot N, van Oijen M, Kessels K, et al. Prediction scores or gastroenterologists' Gut Feeling for triaging patients that present with acute upper gastrointestinal bleeding. *United Eur Gastroenterol J*. 2014;2:197–205.
41. Yaka E, Yilmaz S, Dogan NO, et al. 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.