Author’s response to reviews

Title: AIMS65 scoring system is comparable to Glasgow-Blatchford score or Rockall score for prediction of clinical outcomes for non-variceal upper gastrointestinal bleeding

Authors:

Min Seong Kim (kingtta21@naver.com)
Jeongmin Choi (doct00@hanmail.net)
Won Chang Shin (wcshin@paik.ac)

Version: 1 Date: 27 Oct 2018

Author’s response to reviews:

REVIEWER COMMENTS:

Reviewer reports:
Chikara Iino (Reviewer 1): Comment to the Authors

The form of your manuscript is almost same as recently published literature suggesting the usefulness of the AIMS65 in predicting the mortality for patients with upper gastrointestinal bleeding. Although a certain amount of patients were included in this study, there is no originality and novelty.

1. The research topic of this manuscript has already been performed in several studies. Do authors have any new recommendation for daily practice according to these results?

Response: Although several studies have performed performance of AIMS65, GBS, and other scoring systems, this study is a large scale of study performed in Korea. The causes of UGI bleeding differ among countries. The prevalence of variceal bleeding is higher in Korea than in Western countries. Limited data are available on validation of scoring systems in patients with non-variceal UGI (NVUGI) bleeding in Korea.

Authors showed that the AIMS65 score in NVUGI bleeding patients was comparable to the GBS or Rockall scoring systems when predicting mortality, rebleeding, ICU admission, and endoscopic intervention in Korean patients. The AIMS65 score is easy to calculate using variables routinely available in the emergency clinical situation, and has the advantage that it can be performed before an endoscopy. Therefore, we recommend AIMS65 for prediction of severity of GI bleeding in daily practice.
2. The authors should describe the criteria for the transfusion and ICU admission.

Response: We added the following in the method section.
Transfusion was required if hemoglobin was below 8 g/dL.
ICU admission was considered as following conditions: if the patients have any symptoms of confusion or altered mentality; presence of hemodynamic instability (systolic blood pressure <90 mmHg, need for vasoactive drugs); or severe comorbid illness including heart failure, chronic renal failure, liver cirrhosis, or chronic lung disease. Decision of ICU admission was made finally by ICU unit attending physician.

3. The authors should describe the sentence in methods about excluding variceal bleeding.

Response: We added the following in the method section
Esophageal variceal bleeding or gastric variceal bleeding on endoscopic finding was excluded.

4. In 5 death patients with uncontrolled bleeding, were they treated with the interventional radiology or operational therapy to control bleed?

Response: In the cases of 5 deaths due to uncontrolled bleeding, two patients underwent angiographic embolization and two patients underwent angiography followed by surgery. One patient died of massive hematemesis and hypovolemic shock during endoscopic therapy.

5. Were the patients with Forrest IIa or IIb gastric ulcer treated with endoscopic intervention? Because endoscopic intervention was secondary outcomes, author should describe the definition for endoscopic intervention.

Response: All patients received intravenous proton pump inhibitor infusion before upper endoscopy. Upper endoscopy was performed by endoscopy specialists within the first 24 hours. Endoscopic therapies were applied as following conditions: (1) peptic ulcers with active spurting hemorrhage (type Ia), oozing hemorrhage (type Ib), or a nonbleeding visible vessel (type IIa); (2) dieulafoy’s lesion (3) angiodyplasia (4) any lesions with active bleeding. Adherent blood clot (type IIb), flat spot (type IIc), clean base ulcer (type III), or lesions with no active bleeding stigmata including acute gastric mucosal lesion or Mallory-Weiss tear were initially treated medication without endoscopic therapy. Endoscopic therapies included injection of adrenaline, hemostatic forcep electrocoagulation, argon plasma coagulation, or application of endoscopic clips. In cases of failed endoscopic hemostasis, transarterial embolization was preferred over surgery. When rebleeding occurred after successful endoscopic therapy, endoscopic therapy was initially preferred. Surgery as salvage therapy was performed when embolization was unsuccessful.

6. Rebleeding occurred in 65 patients. Are there any differences in endoscopic findings, medications, or treatment?

Response: All patients received the same medication and the same treatment according to emergency bleeding protocols. Rebleeding occurred in patients with elderly, or chronic kidney disease.
REQUESTED REVISIONS:

Comments for the authors
1. Although you have studied 512 patients, the mortality rates being only 3% limits the full advantages of all the scoring systems for mortality as shown by only 36, 9, and 1 patient with AIMS65 scores of 3, 4, and 5 resp. while the differences between scores of 1, 2, and 3 are not statistically robust at all. This is a major problem with mortality as the main outcome. Have you done any attempt to show the power (statistical robustness) of your overall AIMS65 score for mortality with only 3% risk of death?

Response: We searched the adequate sample size in one article and found 2814 patients were required to show sufficient power for comparative analyses on mortality of 5%. We estimated more than 3000 patients in the case of only 3% risk of death. Our study has limited number of patients due to single center study. We included this in the discussion section. Despite the large number of patients were included in this study, mortality occurred in only 17 patients (3.3%). Because death events were rare, each scoring system to predict mortality was less accurate than previously reported [7]. One study estimated adequate sample size to compare all outcomes, assuming ROC of 0.86 for mortality, 0.78 for AIMS65 and RS with an α of 5%, power of 90%, rank correlation of 0.5, and mortality of 5% [7]. A sample size of 2814 patients was required to show sufficient power for comparative analyses on mortality. Our study showed the AIMS65 score seemed to be superior to the GBS for predicting mortality, but failed to show statistically significant difference between the groups. Limitation of the study is single center study and death events were rare. Multicenter study is needed to obtain sufficient statistical power.

Maybe consider refocusing the study on a more common primary outcome like re-bleeding. Something that might have more relevance to the need for endoscopy. I see that the rebleeding scores as well as the other secondary variables also are not different; this markedly detracts from the support for AIMS65 as being a better more accurate scoring system.

Response: Risk scoring systems were used for prediction of mortality, rebleeding, need for intervention, and transfusion. However, based on our data, prediction values (AUROC) of AIMS65, GBS for mortality were higher than that for other clinical outcomes (AUROCs of AIMS65=0.84, GBS=0.72 for mortality, AIMS65=0.58, GBS=0.55 for rebleeding, AIMS65=0.57, GBS=0.61 for endoscopic intervention). Because AIMS65 or GBS scoring systems derived originally for predicting mortality, we guess AIMS65 or GBS better predicted mortality over rebleeding/endoscopic intervention, or transfusion. That is why we set mortality as primary outcome.

One of the major potential problems of your pairwise comparisons is that it appears that you did not correct for the multiple comparisons by something like a Bonferonni correction- why?

Response: We included it in the discussion section.
“In a Bonferroni correction to adjust for the multiple comparisons in the scoring systems (in which the P values were multiplied by 4 for the comparison with an alpha level of 0.05), the corrected P values were 0.0125 for the combination group versus.”

2. Why should one use the AIMS65 when the combined clinical outcomes of ALL the scores are the
same? You really need to make a much stronger point for supporting AIMS65 over the others. Did you consider adding the results of the EGD to the AIMS65 score to see if that improved your scoring? If not, why not?

Response:
We included it in the discussion section.

The GBS scoring system is limited by weighting, which makes calculation difficult. RS scoring system also has limitations. Weighting leads to complexity in calculation and it requires endoscopic data for calculation, impossible to apply at the time of presentation. Those systems are somewhat difficult to apply in routine clinical practice due to their complexity. AIMS65 has only 5 components (albumin, INR, mental status, systolic BP, age) and each component is the same value of 1 point. Therefore, AIMS65 is much easier to apply in clinical practice.

We believe that adding the results of the EGD to the AIMS65 score would have little impact on prediction value because AIMS65 predicts mortality irrespective of the results of EGD.

3. What does a difference in points mean functionally when comparing scores between the different systems.

Response: AIMS65 scoring system has objective component (albumin, INR, systolic BP, and age) and has no weighting points when comparing GBS or Rockall scoring systems.

Weighting derived from

4. Page 9 line 19 why are they more difficult? This is or should be your entire argument so define why!

Response: In order to be an effective tool for risk stratification, a measure should be easy to use and accurately predict bleeding outcomes. For example, the CHADS2 scoring system, which is used to predict cerebral vascular risk in patients with atrial fibrillation, is widely used because of its accessibility and simplicity. The GBS scoring system is limited by weighting, which makes calculation difficult. RS scoring system also has limitations. Weighting leads to complexity in calculation and it requires endoscopic data for calculation, impossible to apply at the time of presentation. Those systems are somewhat difficult to apply in routine clinical practice due to their complexity.

5. Page 9 lines 28-31. In the Discussion you introduce the term sensitivity and specificity- these MUST be reported in your RESULTS section also you say the sensitivity is good but is it better than the others- indeed when one looks at table 3 the sensitivity looks pretty similar while there may be an advantage to specificity. Why isn't this discussed in the results section?

Response: We added in the statistical analysis section and result section.

In a ROC curve the true positive rate (sensitivity) is plotted in function of the false positive rate (100-specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold.

. Sensitivity of AIMS65 was from 41.5% (CI 19.4-54.4, p <0.001) to 88.2% (CI 63.6-98.5, p <0.0001), which was similar to other scoring systems, ranging from 71.3 (CI 68.1-76.6, p <0.001) to 78.6 (CI 74.1-82.6, p <0.001), respectively.

6. Finally I doubt that many of the readers will know the variables that these other scoring systems use
in their scoring- why not have a table with the variables of each scoring system so that the reader will
not have to try to find the other scoring systems - most readers will not take the time or the effort to
look all the other ones up and they may not have access to a search engine when reading this article.
This is my major suggestion to you- such a table will only help you ( hopefully) to make the argument
that AIMS65 is easier than the other systems because that should be your main argument, not the risk
of mortality because that statistic is much less robust than you make it out to be.

Response: We added the table in the table 4.