

BMJ Open Validation of the ICH score in patients with spontaneous intracerebral haemorrhage admitted to the intensive care unit in Southern Spain

Sonia Rodríguez-Fernández,^{1,2} Encarnación Castillo-Lorente,³ Francisco Guerrero-Lopez,³ David Rodríguez-Rubio,⁴ Eduardo Aguilar-Alonso,⁵ Jesús Lafuente-Baraza,⁴ Francisco Javier Gómez-Jiménez,² Juan Mora-Ordóñez,⁶ Ricardo Rivera-López,⁷ María Dolores Arias-Verdú,⁶ Guillermo Quesada-García,⁶ Miguel Ángel Arráez-Sánchez,⁸ Ricardo Rivera-Fernández⁹

To cite: Rodríguez-Fernández S, Castillo-Lorente E, Guerrero-Lopez F, *et al.* Validation of the ICH score in patients with spontaneous intracerebral haemorrhage admitted to the intensive care unit in Southern Spain. *BMJ Open* 2018;**8**:e021719. doi:10.1136/bmjopen-2018-021719

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-021719>).

Received 18 January 2018
Revised 31 May 2018
Accepted 22 June 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Eduardo Aguilar-Alonso; eaguaia@yahoo.es

ABSTRACT

Objective Validation of the intracerebral haemorrhage (ICH) score in patients with a diagnosis of spontaneous ICH admitted to the intensive care unit (ICU).

Methods A multicentre cohort study was conducted in all consecutive patients with ICH admitted to the ICUs of three hospitals with a neurosurgery department between 2009 and 2012 in Andalusia, Spain. Data collected included ICH, Glasgow Coma Scale (GCS) and Acute Physiology and Chronic Health Evaluation II (APACHE-II) scores. Demographic data, location and volume of haematoma and 30-day mortality rate were also collated.

Results A total of 336 patients were included. 105 of whom underwent surgery. Median (IQR) age: 62 (50–70) years. APACHE-II: 21 (15–26) points, GCS: 7 (4–11) points, ICH score: 2 (2–3) points. 11.1% presented with bilateral mydriasis on admission (mortality rate=100%). Intraventricular haemorrhage was observed in 58.9% of patients. In-hospital mortality was 54.17% while the APACHE-II predicted mortality was 57.22% with a standardised mortality ratio (SMR) of 0.95 (95% CI 0.81 to 1.09) and a Hosmer-Lemeshow test value (H) of 3.62 (no significant statistical difference, n.s.). 30-day mortality was 52.38% compared with the ICH score predicted mortality of 48.79%, SMR: 1.07 (95% CI 0.91 to 1.23), n.s. Mortality was higher than predicted at the lowest scores and lower than predicted in the more severe patients, (H=55.89, $p<0.001$), Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva calibration belt ($p<0.001$). The area under a receiver operating characteristic (ROC) curve was 0.74 (95% CI 0.69 to 0.79).

Conclusions ICH score shows an acceptable discrimination as a tool to predict mortality rates in patients with spontaneous ICH admitted to the ICU, but its calibration is suboptimal.

INTRODUCTION

Patients with spontaneous, non-traumatic acute intracerebral haemorrhage (ICH) represent 9%–27% of all strokes worldwide, with an annual incidence of 10–30 per

Strengths and limitations of this study

- This is a prospective multicentre cohort study with the purpose of validating the intracerebral haemorrhage (ICH) score, including all consecutive patients presenting with diagnosis of spontaneous ICH and admitted to intensive care unit (ICU).
- When predicting mortality, we compared the observed discrimination and calibration for ICH score with the corresponding values analysing Acute Physiology and Chronic Health Evaluation II score in a similar way.
- The statistical analysis was performed through conventional tests used in validation studies, including area under the ROC curve, Hosmer-Lemeshow and Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva tests.
- We did not analyse in this study any clinical differences between patients coming from the emergency department and patients admitted to ICU.
- Considering pupillary changes as a modification item in the ICH score might increase its prognostic accuracy when assessing specifically ICU patients: a bigger number of patients coming from more participating hospitals would increase the reliability of this hypothesis.

100 000 population.^{1–7} The highest mortality rate among patients with stroke is observed in those with ICH, and it is estimated at close to 50%.^{2 8 9} A large proportion of patients with ICH are admitted to the intensive care unit (ICU), and these have a significantly higher mortality rate. Specifically so in hospitals that participate in organ donation programmes, where these patients with ICH are often admitted to the ICU with little or no chances of survival in order to consider organ donation.

Recently, we published a paper about the usefulness of surgery in patients with spontaneous ICH,¹⁰ demonstrating that the mortality rate was particularly high in the population of this multicentre cohort study. Therefore, we posed the question about whether or not this high observed mortality rate was in accord with the predicted one using one of the most widespread prognostic scores (PSs) for this kind of patients: the ICH score.⁹

The ICH score was originally developed for patients with spontaneous ICH admitted to the emergency department (ED) and not to ICU. It is a simple PS that analyses just five variables and provides quick, easy and useful prognostic information for health professionals and relatives in the decision-making process.⁹ As a quality control this PS will allow us to analyse whether the observed mortality in an analysed population is less or greater than the described one by the original PS author and validated with this purpose.^{11 12}

Several specific PSs have been proposed for patients admitted to ICU.^{11 12} Among the most used ones are the Acute Physiology and Chronic Health Evaluation (APACHE), the Simplified Acute Physiologic Score (SAPS) and the Mortality Probability Model (MPM). A previous version of one of these, APACHE-II,¹³ it is probably one of the most used PSs worldwide. In comparison to the ICH score, these ICU prognostic systems analyse a more exhaustive number of variables, providing more accurate prognostic information. However, they are more laborious and difficult to calculate and therefore not so easily applicable in daily clinical practice.

Any given PS should be evaluated in populations of patients different from that in which it was originally described. The original ICH score study was performed in ED patients, and it is mainly in this group where most current studies have been performed. Some research has been performed exclusively in ICU patients,^{14–16} although they have not focused in the validation of the ICH score comparing it with other accurate ICU scores.

As the popular and widespread ICH score is frequently used in current ICU practice, we aimed our research to validate it in the critical care patient group. The objective of this study was the validation of the ICH score in patients with a diagnosis of ICH admitted to the ICU.

MATERIALS AND METHODS

We performed a multicentre cohort study including all consecutive 336 patients presenting with a diagnosis of spontaneous ICH from October 2009 to January 2012 to the ICUs of three hospitals in Andalusia (Regional University Hospital of Málaga, University Hospital Virgen de las Nieves, Granada and Neurotrauma Hospital, Jaén). All participating hospitals are tertiary referral centres for neurosurgical conditions in the region, providing comprehensive facilities for treatment of patients with acute stroke, including neurology, neuroradiology, neurosurgery and critical care. These hospitals frequently

collaborate in multicentre studies^{10 17} and participate in organ donation programmes.

The inclusion criteria were patients over 14 years of age with acute ICH confirmed on CT imaging of the brain with or without intraventricular extension. Patients with a history of trauma, those with aneurysmal subarachnoid haemorrhage or underlying parenchymal lesions were excluded.

After admission patients underwent diagnostic imaging and treatment procedures according to the best clinical practice guidelines at the time of inclusion.¹⁸

The attending neurosurgeon was routinely consulted to assess the need for surgical haematoma evacuation and/or insertion of external ventricular drainage (EVD). The decision for surgical treatment was made following local guidelines for clinical practice or when the attending neurosurgeon considered it necessary. When a surgical decision was made for clot evacuation, a standard craniotomy with corticotomy and haematoma aspiration was carried out in all cases, with or without EVD placement.

For this cohort study, data were collected by specifically trained physicians in each hospital. Data collected at admission included age, admission Glasgow Coma Scale (GCS), APACHE-II, pupillary changes at ICU admission, haematoma volume and location (haematomas originated in the supratentorial compartment were classified into lobar vs basal ganglia/thalamic), presence of intraventricular haemorrhage (IVH) and ICH score. Volume of the haematoma was calculated according to the admission CT scan and applying the formula $A \times B \times C / 2$, being A, B and C the maximal diameters of the clot in the three planes of space.^{8 19} ICH score was used evaluating five items: GCS (13–15:0, 5–12:1, 3–4:2 points), age (<80:0, >80:1 point), infratentorial origin of ICH (No:0, Yes:1 point), ICH volume (<30cc:0, >30cc:1 point) and presence of IVH (No:0, Yes:1 point).⁹ We collected GCS on ICU admission and worst GCS on first ICU day. Although the original ICH score study assessed GCS score at discharge from the emergency room, this value is similar to the GCS on ICU admission used in our study.

Intervention and outcome data were subsequently collected including insertion of EVD and/or surgical evacuation of the clot, length of ICU admission, and 30 days, ICU and overall hospital mortality. Most of these variables were routinely registered in the clinical notes and easily identifiable from the electronic records with a low rate of missing data. Data collection was prospective in Málaga (148 patients) and Granada (93 patients). Jaén Hospital (95 patients) joined the study later (January 2010), so data collection there was partially retrospective. Data were analysed anonymously.

Quantitative variables were expressed as median and IQR, qualitative variables as percentages and frequencies. Non-parametric tests were used to compare continuous variables, applying the Mann-Whitney U test for comparisons of two independent samples. Multivariate analysis was performed using a multiple logistic regression model. We included different variables significantly related to

Table 1 Relationship between hospital mortality and rest of variables (age expressed in years, GCS, APACHE-II and ICH score in points, length of ICU stay in days)

Variables	Total (n=336)	Survivors (n=154)	Non-survivors (n=182)	P values
Age	62(50–70)	58(45–66)	65(55–74)	<0.001
Admission GCS	7 (4–11)	10 (7–14)	6 (4–8)	<0.001
Worst GCS on first ICU day	6 (3–10)	8.5 (6–13)	4 (3–7)	<0.001
Length of ICU stay	5 (2–10)	7 (3–14)	3 (1–7)	<0.001
APACHE-II	21 (15–26)	17 (12–22)	24 (20–29)	<0.001
Probability of death (APACHE-II)	0.59 (0.35–0.79)	0.39 (0.26–0.6)	0.74 (0.54–0.85)	<0.001
ICH score	2 (2–3)	2 (1–3)	3 (2–4)	<0.001
Probability of death (ICH)	0.26 (0.26–0.72)	0.26 (0.13–0.60)	0.72 (0.26–0.97)	<0.001
Admission pupillary changes*				<0.001
Normal	71.8% (n=239)	91.5% (n=140)	55% (n=99)	
Anisocoria	17.1% (n=57)	8.5% (n=13)	24.4% (n=44)	
Bilateral non-reactive mydriasis	11.1% (n=37)	0% (n=0)	20.6% (n=37)	
Volume of haematoma				0.031
<30 mL	44.9% (n=151)	51.3% (n=79)	39.6% (n=72)	
>30 mL	55.1% (n=185)	48.7% (n=75)	60.4% (n=110)	
Origin of the ICH				0.514
Supratentorial	78.3% (n=263)	79.9% (n=123)	76.9% (n=140)	
Infratentorial	21.7% (n=73)	20.1% (n=31)	23.1% (n=42)	
Haematoma site in supratentorial location†				0.6
Lobar	54% (n=141)	55.7% (n=68)	52.5% (n=73)	
Basal ganglia/thalamic	66% (n=120)	44.3% (n=54)	47.5% (n=66)	
Intraventricular haemorrhage				0.03
No	41.1% (n=138)	50% (n=77)	33.5% (n=61)	
Yes	58.9% (n=198)	50% (n=77)	66.5% (n=121)	
External ventricular drain				0.14
No	79.5% (n=267)	76% (n=117)	82.4% (n=150)	
Yes	20.5% (n=69)	24% (n=37)	17.6% (n=32)	
Surgical treatment				<0.001
No	68.8% (n=231)	59.4% (n=95)	77.3% (n=136)	
Yes	31.2% (n=105)	40.6% (n=65)	22.7% (n=40)	

Quantitative variables are expressed as median and 25th–75th percentiles interval.

*Admission pupillary changes data were missed in three patients.

†Haematoma site in supratentorial location data were missed in 2 of 263 patients with supratentorial origin of the haematoma. Calculations were made regarding these 261 patients.

APACHE-II, Acute Physiology and Chronic Health Evaluation II; GCS, Glasgow Coma Scale; ICH, intracerebral haemorrhage; ICU, intensive care unit.

mortality (table 1) in this multivariate analysis, considering that they could improve the prognostic accuracy of the ICH score without increasing its complexity.

Standard mortality ratio (SMR), Hosmer-Lemeshow²⁰ test and (Gruppo Italiano per la Valutazione degli Interventi in Terapia Intensiva, GiViTi) calibration belt^{21 22} were used to assess the agreement between observed and predicted mortality. The area under the ROC curve was used to analyse discrimination.²³ PSpP (psppire.exe 0.7.8.) and 'R V.3.4.1' were used for statistical analysis.

Patient and public involvement

The patients and or public were not involved in the design, in the recruitment or conduct of the study.

RESULTS

Three hundred and thirty-six patients were recruited: 148 from Málaga, 93 from Granada and 95 from Jaén hospital. Median (IQR) age was 62 (50–70) years and GCS on ICU admission was 7 (4–11) points. On admission median

Table 2 Performance of the APACHE-II score: Goodness of fit of general APACHE-II model using Hosmer-Lemeshow χ^2 statistic test

Probability of death (a)	Cases, n	Deaths, n		Survivors, n	
		Observed	Predicted	Observed	Predicted
≤0.20	19	1	2.55	18	16.45
>0.20–0.40	75	19	22.22	56	52.78
>0.40–0.60	64	29	31.01	35	32.99
>0.60–0.80	79	57	56.18	22	22.82
>0.80	78	65	68.61	13	9.39

The Hosmer-Lemeshow test was 3.62 (not statistically significant). APACHE-II, Acute Physiology and Chronic Health Evaluation II.

APACHE-II severity score was 21 (15–26) points and ICH score was 2 (2–3) points. 28.2% of the patients had pupillary anomalies on admission (11.1% had bilateral non-reactive mydriasis). In 263 cases, the origin of the ICH was supratentorial and IVH was present in 58.9% of the patients. One hundred and five patients were surgically treated.

Hospital mortality rate was 54.17%. Patients who died during their hospital stay were significantly older, had lower GCS on admission and higher APACHE-II scores (table 1). Predicted mortality rate according to APACHE-II was 57.22%, and it was comparable with the observed one. SMR was 0.95 (95% CI 0.81 to 1.09), with no significant statistical difference (n.s.). The agreement between the APACHE-II predicted and observed mortality rates was also analysed using the Hosmer-Lemeshow test (H=3.62, n.s., table 2, figure 1) and the GiViTi calibration belt (p=0.430, n.s., supplemental digital content—figure 2), with no statistically significant difference observed. Discrimination for the APACHE-II using the area under the ROC curve was 0.80 (95% CI 0.74 to 0.84).

The observed 30-day mortality rate was 52.38% compared with a predicted ICH score mortality rate of 48.79%. SMR was 1.07 (95% CI 0.91 to 1.23), n.s. Observed and predicted mortality rates in the different ICH score groups (table 3) were: 0 points (n=10): 10% vs 0%; 1 point (n=54): 22.2% vs 13%; 2 points (n=115): 42.6% vs 26%;

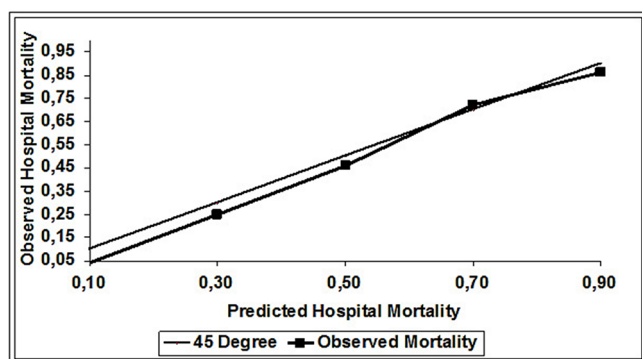


Figure 1 Predicted versus observed hospital mortality for APACHE-II model. APACHE-II, Acute Physiology and Chronic Health Evaluation II.

3 points (n=102): 65.7% vs 72%; 4 points (n=47): 85.1% vs 97%; 5 points (n=7): 85.71% vs 100%; 6 points (n=1): 100% vs 100%. The observed mortality rate was therefore higher in comparison to that predicted in the low ICH scores, and lower in the high scores (3, 4 and 5 points). Hosmer-Lemeshow test (H) was 55.89 (p<0.001), with significant statistical difference in the agreement between the ICH score predicted and observed mortality rates (table 4). The GiViTi calibration belt showed a significant difference (p<0.001) between both rates as well (supplemental digital content—figure 3). Discrimination for the ICH score using the area under the ROC curve was 0.74 (95% CI 0.69 to 0.79).

28.2% of patients (n=94) showed pupillary anomalies (anisocoria or bilateral mydriasis) on admission, with a 30-day observed mortality rate of 85% and an ICH score predicted one of 68.5%. When analysing the particular subgroup of 37 patients (11.1%) with bilateral mydriasis, 30-day mortality was 100% vs 78.11% ICH score predicted rate. In 239 patients with no pupillary anomalies, 30-day observed mortality rate was 39.33%, very similar to the 42.89% predicted by the ICH score.

Multiple logistic regression analysis showed a significant relationship between 30-day hospital mortality rate and ICH score predicted mortality, OR 1.02 (95% CI 1.01 to 1.03), and with pupillary anomalies seen on admission to ICU, OR 5.72 (95% CI 2.96 to 11.04). The area under the ROC curve was 0.80 (95% CI 0.75 to 0.84) according to this model. When including in the model the variable worst GCS on first ICU day, for ICH score the OR was 1.59 (95% CI 1.17 to 2.16), for pupillary anomalies OR was 3.90 (95% CI 1.97 to 7.75) and for worst GCS on first ICU day the OR was 0.83 (95% CI 0.76 to 0.91). The area under the ROC curve was 0.81 (95% CI 0.77 to 0.86) according to this model.

DISCUSSION

Our study shows that ICH score is useful for evaluating the severity of patients with ICH admitted to ICU, although it does not meet all the necessary requirements for its validation.

Its discrimination is acceptable, but with respect to its calibration we have detected a significant difference between the observed and predicted mortality rates, this difference being slightly higher than the accepted one for validation. This difference does not seem to be important enough to invalidate this predictive scale, and therefore the ICH score can be considered as a useful tool for the evaluation of patients with spontaneous ICH. An adaptation to the ICU patients may be considered in the future to achieve a more precise predictive value of the score in the critical patient.

Our research also shows that patients with pupillary anomalies on admission are a very important source of discrepancies between observed and predicted mortality. The three participating hospitals included in our analysis are centres with organ donation programmes, and that

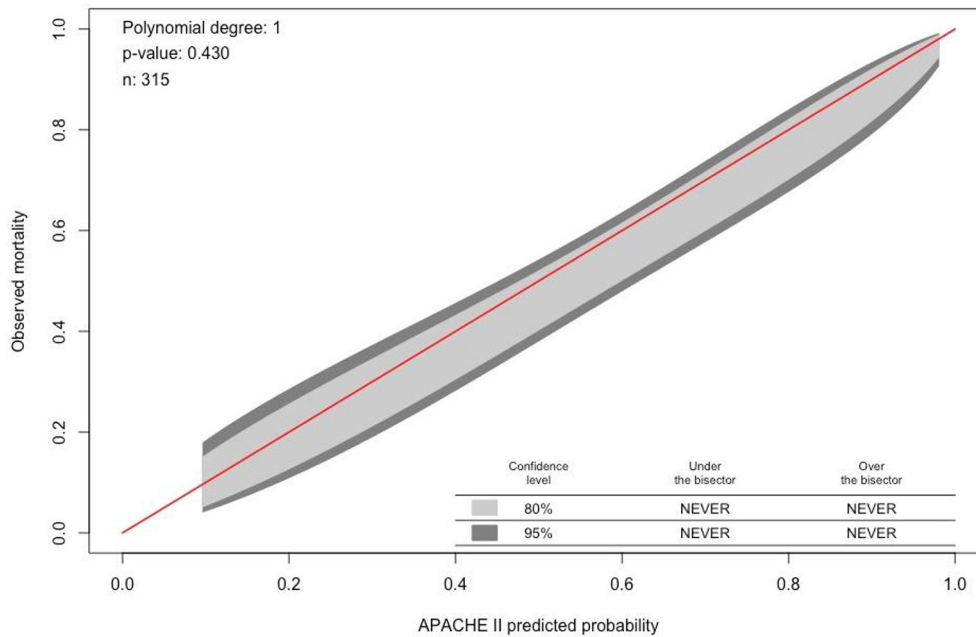


Figure 2 APACHE-II calibration. APACHE-II, Acute Physiology and Chronic Health Evaluation II.

increases the number of observed patients with pupillary anomalies (including those with bilateral mydriasis). This fact can be extended to almost all ICUs with Donation and Transplant Programme. Multivariate analysis in our research showed that the presence of pupillary anomalies (classified in anisocoria and bilateral non-reactive mydriasis) it is a predictive factor that complements the ICH score, and it could be used as a possible adaptation of it. Pupillary changes are routinely examined in patients with spontaneous ICH, and further research could clarify if this easily identifiable sign would improve calibration of the ICH score when assessing specifically ICU patients.

Our study also shows that worst GCS on first ICU day (an included item in APACHE-II score) is a significant factor after multivariate analysis, this sequential clinical variable could also contribute to a more accurate prognosis.

ICH score has been validated worldwide through several studies^{9 24–27} in different populations.^{28–32} However, the ICH score has been scarcely validated in ICH patients specifically admitted to the ICU.^{14–16} This prognostic tool was originally created for evaluation of ICH patients

admitted to the ED before they were moved to ward, ICU or operating theatre, or excluded from resuscitation manoeuvres.

We have not specifically studied differences between ICU and ED patients, neither have we collected data about Do Not Resuscitate/Do Not Intubate orders in the ED. Some patients never even reach the ICU from the ED, and as a result, are excluded from the ICU validation and not the ED. ICU admitted patients usually do not include those with mild symptoms and minimal haemorrhage, but instead those coming from the operating theatre after surgical evacuation of the clot and sometimes those considered non-recoverable and admitted just for support measurements/transplant intention. These clinical biases introduced by a different type of population in comparison to ED patients can explain the suboptimal calibration of ICH score when applied exclusively in ICU patients, and establishes a limitation in our study that will need clarification in further research. There may be characteristics in patients with ICH that may explain why these patients survived to the ICU when others did not.

Table 3 Observed and predicted 30-day mortality rates according to intracerebral haemorrhage (ICH) score

ICH score	n	Observed (%)	Predicted (%)
0	10	10	0
1	54	22.2	13
2	115	42.6	26
3	102	65.72	72
4	47	85	97
5	7	85.71	100
6	1	100	100

Table 4 Performance of the ICH score: Goodness of fit of general ICH score model using Hosmer-Lemeshow χ^2 statistic test

Probability of death (a)	Cases, n	Deaths, n		Survivors, n	
		Observed	Predicted	Observed	Predicted
≤0.25	64	13	7.02	51	56.98
>0.25–0.5	115	49	29.9	66	85.1
>0.5–0.75	102	67	73.4	35	28.6
>0.75	55	47	53.59	8	1.41

The Hosmer-Lemeshow test was 55.89 (p<0.001).
ICH, intracerebral haemorrhage.

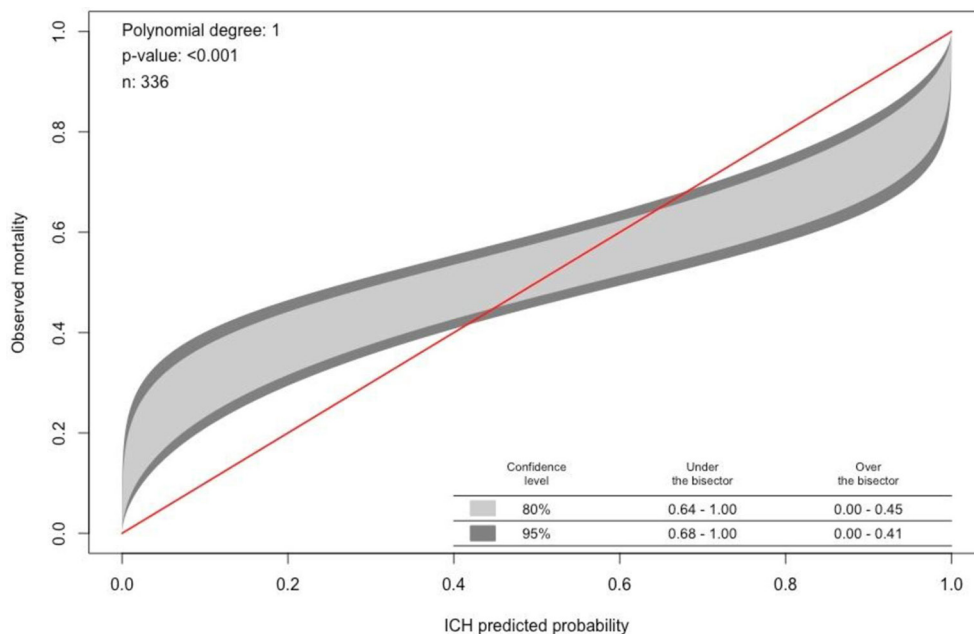


Figure 3 ICH calibration. ICH, intracerebral haemorrhage.

The rate of operated patients (31.2%) in our study can be considered high, but similar to other series.¹⁶ The role of surgery for most patients with spontaneous ICH remains controversial. Randomised trials comparing surgery to conservative management have not demonstrated a clear benefit for surgical intervention. Moreover, the generalisability of the results of these trials can be questioned, because patients at risk for herniation were likely excluded and the largest and most recent studies had high rates of treatment group crossover from conservative management to surgery. STICH trial³³ found no overall statistically significant difference in mortality or functional outcome between treatment groups. Notably, 26% of patients initially assigned to conservative management ultimately underwent surgery. Our data were collected between the publication of the results of the STICH³³ and STICH-II trials,³⁴ which could have influenced the neurosurgical decision-making in the three participating centres.

These studies performed exclusively on ICU patients have not focused on the topic of the validation of the score in critically ill patients. Fang *et al*¹⁶ analysed a group of ICU patients with non-traumatic ICH, focusing on the influence of coagulation disorders in these patients. The authors also provide data on the number of patients included in each of the scores and their mortalities, which allows us to see the functioning of the ICH score in their patients; in the 0 points group (n=15) the observed and predicted mortality rate was 0%, in the 1 point group (n=22) the observed and predicted rates were 9% and 13%, respectively, in the 2 points group (n=43) they were 23% vs 26%, in the 3 points group (n=19) 58% vs 72%, and in the 4 points group (n=10) the observed and predicted mortality rates were 50% vs 97%. Thirtydays observed and predicted mortality rates were 25.7% vs

34.23%, respectively. Performing the Hosmer-Lemeshow test on these data, we find a statistically significant difference ($H=78.23$, $p<0.001$) in the agreement between the observed and predicted mortality rates. As we noticed in our patients, the most significant differences are found in the more severe groups of patients, where ICH score overestimates mortality rates. Another study specifically performed in ICU patients is that of Godoy *et al*,¹⁴ where observed and predicted mortality rates according to ICH scores were: 0 points: 0% vs 0%, 1 point: 2.9% vs 13%, 2 points: 30.8% vs 26%, 3 points: 61.1% vs 72%, 4 points: 88.2% vs 97%). In this paper, we observe scant discrepancies in the different scores and the authors conclude that there is a good agreement between what was observed and predicted. In a third study performed in ICU patients,¹⁵ similar differences were also noticed in observed and predicted mortalities (0 points: 11.11% vs 0%, 1 point: 12.5% vs 13%, 2 points: 44.44% vs 26%, 3 points: 70% vs 72%, 4 points: 100% vs 97%). In this last study, the number of included patients studied was only 37, and therefore, it is difficult to draw reliable conclusions from it.

The analysis of the data in these papers agrees with the results from our study: the calibration of the ICH score is suboptimal. Significantly, the observed mortalities in the more severe groups of patients are inferior to those predicted by the ICH score. However, we believe that this is an easy-to-use instrument and that it is useful when evaluating the severity of patients to provide the attending doctor and relatives with quick and simple information about a patient's condition and approximate prognosis. The ICH score, however, does demonstrate an inadequate calibration in the most critical patients, and therefore, currently it cannot be considered as a precise outcome predictor.

The calibration discrepancies we observe may warrant an adaptation of the instrument to critical care patients in the future. Several studies have attempted to improve its sensitivity, specificity and its usefulness for the prognostic evaluation of morbidity by adding variations to it. Cheung and Zou²⁶ use a modified ICH score where National Institute of Health Stroke Scale, temperature and pulse pressure are added. Wang *et al*²⁷ and Ruiz-Sandoval *et al*³¹ compared ICH score and ICH grading scale score. Both indexes analyse the same parameters but they attribute different scores to each parameter.

Multiple PSs have been developed for ICU patients. APACHE, SAPS and MPM systems^{12 13} are probably the most widespread ones. These systems can be used on most ICU patients including those with acute ICH. They evidence a discrimination superior to the found one in the ICH score, but they also show the disadvantage of its greater complexity.

PSs resulting from a bigger number of variables have better discrimination. Validation studies of these scores usually do not find discrimination problems, but nevertheless, they show calibration issues, with statistically significant differences between the predicted mortality calculated by the system and the observed one in the study population. We have performed several of these validation studies in which we found good discrimination but not an adequate calibration in different PSs.^{17 35 36}

We analysed in this study the performance of the APACHE-II score as well, showing a higher discrimination than the calculated one for the ICH score. We also found a remarkable good calibration for the APACHE-II, a PS that was created more than 30 years ago, and in which it would be theoretically expected that the score calibration deteriorated over time, due to changes in the type of patients and to improvement in treatments. No significant difference between predicted and observed mortality using APACHE-II was observed with both calibration statistical tests. Current published data on the good functioning of APACHE-II in these type of patients are available.^{37 38}

The ideal PS to be promoted and developed should be simple, with few component variables codified in a simple way and with good discrimination capacity. It is about finding a balance between simplicity and precision when predicting the results.

ICH score was specifically designed for patients with non-traumatic ICH. As described in the material and methods section, it collects the information according to five variables previously identified as significantly related to mortality, and it categorises the continuous variables to make its use simpler. Scores vary from 0 to 6. Its simplicity makes this system intuitive and easy, allowing the physician quickly to quantify severity and chances of survival in the ED. It is a useful method that estimates the chances of recovery in a patient prior to admission in the ICU. It is not a PS specifically designed for ICU patients, but its simplicity makes its use currently widespread among critically ill patient carers. We cannot validate it in terms of

calibration, but its discrimination looks good enough to allow its use in ICU daily clinical practice. The ICH score was developed more than 15 years ago. The advances in diagnosis and treatment of patients with ICH have improved their survival, and that could make the ICH score predictions obsolete. On the other hand, some may claim that advances that significantly influence over mortality in these patients during last years have been scarce, and therefore, ICH score predictions would be still reliably up to date.^{39 40}

Regarding the limitations of our study, we believe that a bigger number of patients coming from more participating hospitals would increase its reliability. However, the number of included patients in our multicentre study is sufficient to draw statistically significant conclusions.

CONCLUSIONS

ICH score shows an acceptable discrimination as a tool to predict mortality rates in patients with spontaneous ICH admitted to the ICU, but its calibration is suboptimal. These calibration discrepancies we observe may warrant an adaptation of the instrument to critical patients in the future.

Author affiliations

- ¹Intensive Care Medicine, Hospital de la Serranía, Ronda, Spain
- ²Programa de Doctorado, Universidad de Granada, Granada, Spain
- ³Intensive Care Medicine, Hospital Virgen de las Nieves, Granada, Spain
- ⁴Neurosurgery Department, Hospital del Mar, Barcelona, Spain
- ⁵Intensive Care Medicine, Hospital Infanta Margarita, Cabra, Spain
- ⁶Intensive Care Medicine, Hospital Carlos Haya, Málaga, Spain
- ⁷Cardiology Department, Hospital Virgen de las Nieves, Granada, Spain
- ⁸Neurosurgery Department, Hospital Carlos Haya, Málaga, Spain
- ⁹Intensive Care Medicine, Complejo Hospitalario de Jaén, Jaén, Spain

Acknowledgements The authors would like to thank to Dr Emilio Sanchez-Cantalejo, MD, (Professor of Statistics and Epidemiology at the Andalusian School of Public Health—Granada, Spain, CIBER Epidemiology and Public Health CIBERESP, Spain) the statistical review and correction of this paper. Mr James Walkden from the Neurosurgery Department in Aberdeen Royal Infirmary also reviewed and corrected the written English. Professor David Mendelow from Newcastle University improved the final version of this paper with his invaluable advice.

Contributors SR-F, EC-L and GQ-G: conception and design of the work. FG-L, JM-O, MDA-V and MAA-S: acquisition data. R-RL and RR-F: analysis and interpretation of data. SR-F, JL-B, EA-A, FJG-J, DR-R and MAA-S: drafting. SR-F, MAA-S and RR-F: final approval of the version to be published.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Ethics approval This study protocol is approved by the Ethical Committee of the three participating hospitals.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The authors declare that there are no additional unpublished data. The data are accessible by requesting the correspondence author.

Author note The results of this article are part of the doctoral thesis: "Validation of the ICH score in patients with spontaneous intracerebral haemorrhage admitted to the Intensive Care unit", University of Granada, Spain.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) 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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Feigin VL, Forouzanfar MH, Krishnamurthi R, *et al*. Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. *Lancet* 2014;383:245–55.
- Labovitz DL, Halim A, Boden-Albala B, *et al*. The incidence of deep and lobar intracerebral hemorrhage in whites, blacks, and Hispanics. *Neurology* 2005;65:518–22.
- Sudlow CL, Warlow CP. Comparable studies of the incidence of stroke and its pathological types: results from an international collaboration. International Stroke Incidence Collaboration. *Stroke* 1997;28:491–9.
- Qureshi AI, Mendelow AD, Hanley DF. Intracerebral haemorrhage. *Lancet* 2009;373:1632–44.
- Zahuranec DB, Gonzales NR, Brown DL, *et al*. Presentation of intracerebral haemorrhage in a community. *J Neurol Neurosurg Psychiatry* 2006;77:340–4.
- Flaherty ML, Woo D, Haverbusch M, *et al*. Racial variations in location and risk of intracerebral hemorrhage. *Stroke* 2005;36:934–7.
- Sacco S, Marini C, Toni D, *et al*. Incidence and 10-year survival of intracerebral hemorrhage in a population-based registry. *Stroke* 2009;40:394–9.
- Broderick JP, Brott TG, Duldner JE, *et al*. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. *Stroke* 1993;24:987–93.
- Hemphill JC, Bonovich DC, Besmertis L, *et al*. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke* 2001;32:891–7.
- Rivera-Fernández R, Guerrero-López F, Rodríguez-Rubio D, *et al*. Survival analysis of surgically evacuated supratentorial spontaneous intracerebral hemorrhage with intraventricular extension. *Neurocirugia* 2016;27:220–8.
- Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. *Crit Care* 2010;14:207.
- Moreno RP. Outcome prediction in intensive care: why we need to reinvent the wheel. *Curr Opin Crit Care* 2008;14:483–4.
- Knaus WA, Draper EA, Wagner DP, *et al*. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818–29.
- Godoy DA, Piñero G, Di Napoli M. Predicting mortality in spontaneous intracerebral hemorrhage: can modification to original score improve the prediction? *Stroke* 2006;37:1038–44.
- Patriota GC, Silva-Júnior JM, Barcellos AC, *et al*. Determining ICH Score: can we go beyond? *Arq Neuropsiquiatr* 2009;67:605–8.
- Fang HY, Lin CY, Ko WJ. Hematology and coagulation parameters predict outcome in Taiwanese patients with spontaneous intracerebral hemorrhage. *Eur J Neurol* 2005;12:226–32.
- López-Caler C, García-Delgado M, Carpio-Sanz J, *et al*. External validation of the Simplified Acute Physiology Score (SAPS) 3 in Spain. *Med Intensiva* 2014;38:288–96.
- Morgenstern LB, Hemphill JC, Anderson C, *et al*. Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2010;41:2108–29.
- Gebel JM, Sila CA, Sloan MA, *et al*. Comparison of the ABC/2 estimation technique to computer-assisted volumetric analysis of intraparenchymal and subdural hematomas complicating the GUSTO-1 trial. *Stroke* 1998;29:1799–801.
- Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92–106.
- Nattino G, Finazzi S, Bertolini G. A new test and graphical tool to assess the goodness of fit of logistic regression models. *Stat Med* 2016;35:709–20.
- Poole D, Carrara G, Bertolini G. Intensive care medicine in 2050: statistical tools for development of prognostic models (why clinicians should not be ignored). *Intensive Care Med* 2017;43:1403–6.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
- Clarke JL, Johnston SC, Farrant M, *et al*. External validation of the ICH score. *Neurocrit Care* 2004;1:53–60.
- Jamora RD, Kishi-Generao EM, Bitanga ES, *et al*. The ICH score: predicting mortality and functional outcome in an Asian population. *Stroke* 2003;34:6–7.
- Cheung RT, Zou LY. Use of the original, modified, or new intracerebral hemorrhage score to predict mortality and morbidity after intracerebral hemorrhage. *Stroke* 2003;34:1717–22.
- Wang W, Lu J, Wang C, *et al*. Prognostic value of ICH score and ICH-GS score in Chinese intracerebral hemorrhage patients: analysis from the China National Stroke Registry (CNSR). *PLoS One* 2013;8:e77421.
- Fernandes H, Gregson BA, Siddique MS, *et al*. Testing the ICH score. *Stroke* 2002;33:1445–6.
- Parry-Jones AR, Abid KA, Di Napoli M, *et al*. Accuracy and clinical usefulness of intracerebral hemorrhage grading scores: a direct comparison in a UK population. *Stroke* 2013;44:1840–5.
- Satopää J, Mustanoja S, Meretoja A, *et al*. Comparison of all 19 published prognostic scores for intracerebral hemorrhage. *J Neurol Sci* 2017;379:103–8.
- Ruiz-Sandoval JL, Chiquete E, Romero-Vargas S, *et al*. Grading scale for prediction of outcome in primary intracerebral hemorrhages. *Stroke* 2007;38:1641–4.
- Meyer DM, Begtrup K, Grotta JC. Is the ICH score a valid predictor of mortality in intracerebral hemorrhage? *J Am Assoc Nurse Pract* 2015;27:351–5.
- Mendelow AD, Gregson BA, Fernandes HM, *et al*. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. *Lancet* 2005;365:387–97.
- Mendelow AD, Gregson BA, Rowan EN, *et al*. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial lobar intracerebral haematomas (STICH II): a randomised trial. *Lancet* 2013;382:397–408.
- Vázquez Mata G, del Mar Jiménez Quintana M, Rivera Fernández R, *et al*. [Severity assessment by APACHE III system in Spain]. *Med Clin* 2001;117:446–51.
- Banderas-Bravo ME, Arias-Verdú MD, Macías-Guarasa I, *et al*. Patients Admitted to Three Spanish Intensive Care Units for Poisoning: Type of Poisoning, Mortality, and Functioning of Prognostic Scores Commonly Used. *Biomed Res Int* 2017;2017:1–9.
- Moon BH, Park SK, Jang DK, *et al*. Use of APACHE II and SAPS II to predict mortality for hemorrhagic and ischemic stroke patients. *J Clin Neurosci* 2015;22:111–5.
- Huang Y, Chen J, Zhong S, *et al*. Role of APACHE II scoring system in the prediction of severity and outcome of acute intracerebral hemorrhage. *Int J Neurosci* 2016;126:1020–4.
- Barnes B, Hanley DF, Carhuapoma JR. Minimally invasive surgery for intracerebral haemorrhage. *Curr Opin Crit Care* 2014;20:148–52.
- Adeoye O, Woo D, Haverbusch M, *et al*. Eligibility for the surgical trial in intracerebral hemorrhage II study in a population-based cohort. *Neurocrit Care* 2008;9:237–41.