BMJ Open

Methodological quality of multivariate prognostic models for intracranial haemorrhages in intensive care units: a systematic review

Jeanne Simon-Pimmel ,[1] Yohann Foucher,[1,2] Maxime Léger,[1,3] Fanny Feuillet,[1,4] Laetitia Bodet-Contentin,[1,5] Raphaël Cinotti,[6] Denis Frasca,[1,7] Etienne Dantan [1]

To cite: Simon-Pimmel J, Foucher Y, Léger M, *et al*. Methodological quality of multivariate prognostic models for intracranial haemorrhages in intensive care units: a systematic review. *BMJ Open* 2021;**11**:e047279. doi:10.1136/bmjopen-2020-047279

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-047279).

Received 24 November 2020
Accepted 25 July 2021

For numbered affiliations see end of article.

Correspondence to
Jeanne Simon-Pimmel;
Jeanne.Pimmel@gmail.com

ABSTRACT

Objectives Patients with severe spontaneous intracranial haemorrhages, managed in intensive care units, face ethical issues regarding the difficulty of anticipating their recovery. Prognostic tools help clinicians in counselling patients and relatives and guide therapeutic decisions. We aimed to methodologically assess prognostic tools for functional outcomes in severe spontaneous intracranial haemorrhages.

Data sources Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses recommendations, we conducted a systematic review querying Medline, Embase, Web of Science, and the Cochrane in January 2020.

Study selection We included development or validation of multivariate prognostic models for severe intracerebral or subarachnoid haemorrhage.

Data extraction We evaluated the articles following the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies and Transparent Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis statements to assess the tools' methodological reporting.

Results Of the 6149 references retrieved, we identified 85 articles eligible. We discarded 43 articles due to the absence of prognostic performance or predictor selection. Among the 42 articles included, 22 did not validate models, 6 developed and validated models and 14 only externally validated models. When adding 11 articles comparing developed models to existing ones, 25 articles externally validated models. We identified methodological pitfalls, notably the lack of adequate validations or insufficient performance levels. We finally retained three scores predicting mortality and unfavourable outcomes: the IntraCerebral Haemorrhages (ICH) score and the max-ICH score for intracerebral haemorrhages, the SubArachnoid Haemorrhage International Trialists score for subarachnoid haemorrhages.

Conclusions Although prognostic studies on intracranial haemorrhages abound in the literature, they lack methodological robustness or show incomplete reporting. Rather than developing new scores, future authors should focus on externally validating and updating existing scores with large and recent cohorts.

Strengths and limitations of this study

- This is the first systematic review of the methodological quality of prognostic tools for severe spontaneous intracranial haemorrhages managed in intensive care units.
- A robust search strategy with no language restriction was performed, leading to a high number of eligible articles.
- This systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement, and we evaluated the articles following the Transparent Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis statement to assess the tools' methodological reporting and pitfalls.
- This systematic review concerns two types of lesions intracerebral haemorrhages and subarachnoid haemorrhages that present different pathophysiologies and clinical courses but similar long-term consequences, leading us to suspect shared methodological issues.
- We were not able to perform a meta-analysis due to the heterogeneity in the included models.

INTRODUCTION

Severe spontaneous intracranial haemorrhages, managed in intensive care units (ICUs), are at high risk of developing complications such as rebleeding or cerebral ischaemia,[1 2] leading to high morbidity and mortality. Intracerebral haemorrhages (ICH) have a mortality rate of 40% at 1 month,[3] while subarachnoid haemorrhages (SAH) have a mortality rate of 25% at 10 years.[4] Survivors have a high rate of vegetative state or severe disabilities.[5] This serious statement highlights the initial issues specific to severe strokes and the challenge physicians and surrogates face in deciding to continue invasive care.[6 7] Indeed, the question arises as to whether advanced resuscitation is justified when the future appears unfavourable.[8]

When considering a limitation of care, the essential issue is to prevent inaccurate self-fulfilling prophecies by predicting outcomes reliably.[9] In such settings, an individual's patient prognostic may be difficult to assess because of the multiplicity of risk factors involved in the evolution of severe intracranial haemorrhages. Multivariable prognostic scores could assist clinicians in counselling patients and relatives and guide therapeutic decisions.

Previous reviews of prognostic tools,[10–14] popular in the field of neurocritical care, have not focused on injuries managed in ICUs, for whom the issue of advanced care pursuits is a concern. Indeed, scores are reliable when validated in the population of interest. They also did not address the methodological quality of the selected articles. The PROgnosis RESearch Strategy (PROGRESS) group recently proposed a framework for prognosis concerns[15 16] that led to the Transparent Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.[17] These recommendations efficiently summarised the process for developing and validating a prognostic scoring system.

The objective of our systematic review was to assess the methodology of existing prognostic tools of functional outcomes in patients with severe spontaneous intracranial haemorrhage managed in ICUs. We chose to conduct this systematic review for the two types of lesions (ICH and SAH). While their pathophysiologies and clinical courses are different, the consequences for long-term functional outcomes are similar. The questions that arise at the beginning of the ICU stay about patients' future and the complex ethical decisions are similar. While prognostic models may differ, the way to develop them should follow a similar modelling process. We suspected that studies presenting prognostic tools share the same methodological issues.

MATERIALS AND METHODS

Search strategy

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (online supplemental table S1).[18] We searched Medline, Embase, Web of Science and the Cochrane databases on 7 December 2017 and updated on 14 January 2020, without date restriction. We used a query based on Medical Subject Heading terms and keywords. Online supplemental file S2 outlines the detailed search strategy.

Study selection

We included all-language studies focusing on adults with severe spontaneous intracranial haemorrhage (ICH or SAH) managed in ICU, or specified explicitly as 'severe' or 'high grade' injury. We did not include criteria on the location, the cause of the haemorrhage or the type of cases (primary or secondary haemorrhage). We did not include paediatric studies or studies uniquely concerning traumatic injuries. We searched for the development and/or validation of prognostic models, predicting outcomes using variables collected before or at the beginning of their ICU stay. The targeted outcomes were mortality, functional outcomes or quality-of-life-related outcomes from ICU-discharge or hospital-discharge through to long-term outcomes. Our non-inclusion criteria were reviews or meta-analyses, full texts not found or conference abstracts, models developed without predictor selection, univariate models or the lack of reported prognostic performance. One reviewer (JS-P) screened references by title and abstract. The full eligible texts were assessed independently by four pairs of reviewers (YF–ML, FF–RC, DF–LB-C and JS-P–ED) and discussions resolved any discrepancies.

Data extraction

We predefined a standardised form for data extraction and evaluation of the risk of bias (online supplemental tables S3 and S4). For each eligible article, we collected the author's name, year and journal, data source and study design, inclusion and exclusion criteria, sample size, population characteristics, predicted outcomes (mortality, functional outcomes and quality-of-life), prediction time (ie, the time when one calculates the prediction), horizon time (ie, the end of the prediction time window), predictive tools, development details (such as variables of the scoring systems), internal validation details, external validation details, missing data information and open comments regarding bias and limitations.

Articles and prognostic tools selection based on quality assessment

To include the articles, we followed the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) [19] and the TRIPOD statements.[20] Specifically, they recommend developing a score from a learning sample set and validating the prognostic performance from an independent sample (internal and/or external validation). This step avoids reporting the prognostic capacities on the training sample only because no internal nor external validation led to overestimating their performance.[21] The articles reporting model development without any validation were thus not retained. They also recommend having a sufficient sample size and a sufficient number of events (known as the effective sample size). We considered at least 250 patients and 50/50 events and non-events as sufficient. The modelling strategy must also consider enough events per predictor, usually at least 10, to avoid overfitting.[22 23] We did not include articles that did not follow these recommendations.

Assessment of the performance of the prognostic tools should use discrimination (ability to differentiate between patients who do or do not experience the event, eg, area under the curve (AUC) receiver operating characteristic (ROC)), calibration (agreement between predictions from the model and observed outcomes) and global measures (simultaneous evaluation of calibration and discrimination, eg, Brier Score). Among the

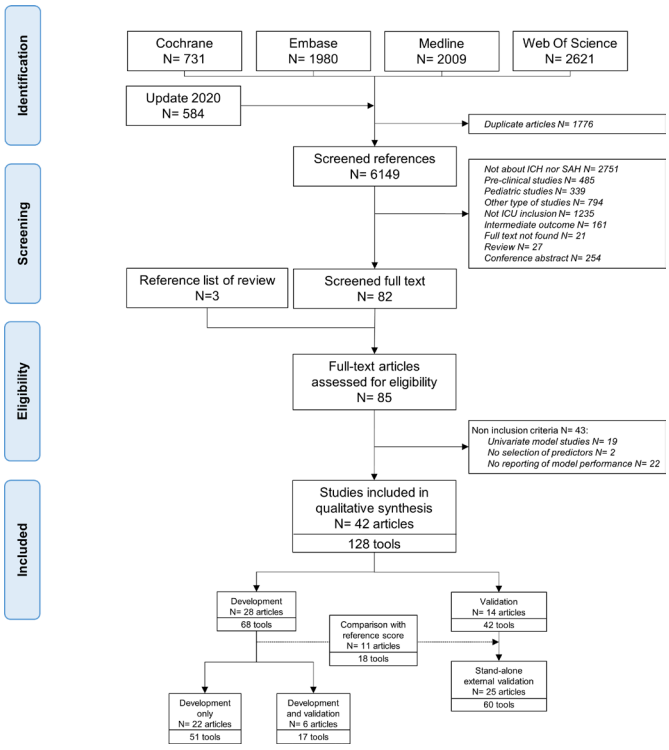

Figure 1 PRISMA flow diagram, selection of included studies. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

included articles, the retained prognostic tools were those presenting good prognostic performances reported on internal and/or external validation.

Patient and public involvement

This study has no patient or public involvement.

RESULTS

Description of studies

The electronic database search identified 6149 unique references. Screening of titles/abstracts and references checking of included articles and reviews identified 85 eligible papers for full-text review. We did not include 43 articles for the following reasons: 19 univariate models, 2 models without predictor selection and 22 multivariate models without performance reporting. Finally, we included 42 articles (figure 1).

All articles were in English. There were 11 articles published before 2010, 12 between 2010 and 2015 and 19 after 2015. The published teams were mainly from Europe (n=17, 40%) and North America (n=14, 33%). Patients were mostly recruited into an ICU (n=33, 79%). Inclusion criteria were heterogeneous in terms of location or aetiology of the haemorrhage. For ICH, most studies included only spontaneous ICH, some excluding malformations and/or coagulation disorder. For SAH, most included aneurysmal SAH online supplemental tables S3 and S4 present the information regarding inclusion and non-inclusion criteria of each study. The pooled mean age was 59.3 years (SD 13.7) (data not available for

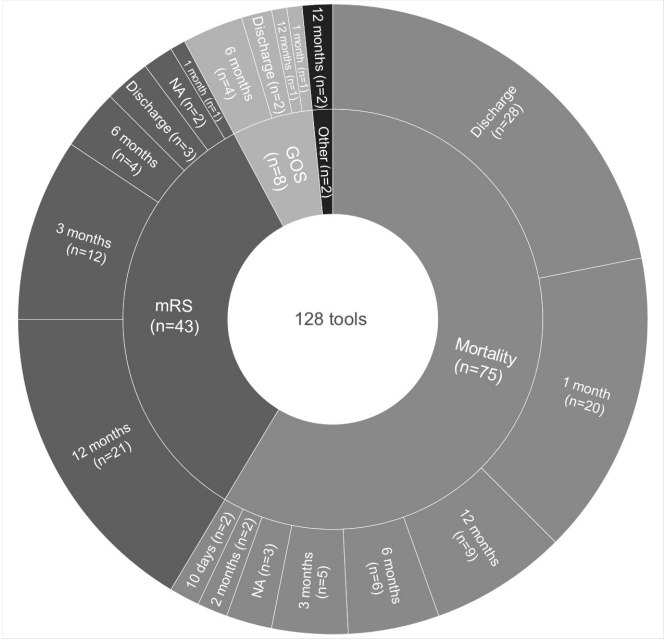

Figure 2 Predicted outcomes and corresponding horizon times of the 128 prognostic tools. GOS: Glasgow Outcome Scale; mRS: modified Rankin Scale; NA, not available.

six studies). Fifty-three per cent (range 21%–73%) were female (missing data for five studies). The 42 eligible articles reported 128 prognostic tools (figure 1): five articles reported one tool, 16 reported two tools, 7 reported three tools and 14 articles more than three tools, differing by their predictors, their types of outcome or their horizon times. Regardless of the types of predicted outcomes, the sample sizes ranged from 68 to 1629 patients (median 290, IQR 128–413), and the number of events ranged from 21 to 786 (median 64.5, IQR 34–164). Regardless of the time of prediction, most of the prognostic tools predicted mortality (n=75, 59%) (figure 2). Fifty-one (40%) tools studied functional outcomes using the modified Rankin Scale (mRS) or the Glasgow Outcome Scale (GOS). The horizon time for mortality data was mostly short-term (67% at discharge or 1 month), unlike functional outcomes (14% at discharge or 1 month) (figure 2). One study predicted the cognitive status and physical quality-of-life at 12 months. The 452 predictors of these 128 tools mainly involved baseline characteristics (n=95, 21%), admission clinical variables (n=104, 23%), biological measures (n=86, 19%), CT variables (n=95, 21%), ICU-evolution variables (n=29, 6%), existing scores (n=40, 9%) and others (n=3, 1%). Most variables were available on admission, others within 72 hours after ICU admission, and few were available throughout the ICU stay. The prediction time was sometimes unknown.

Model development studies

Twenty-eight studies developed prediction models. Online supplemental table S3 provides complete standardised form and references. Twelve articles focused on patients with ICH only, 15 on patients with SAH only and one on patients with both ICH and SAH. Of the 16 articles on

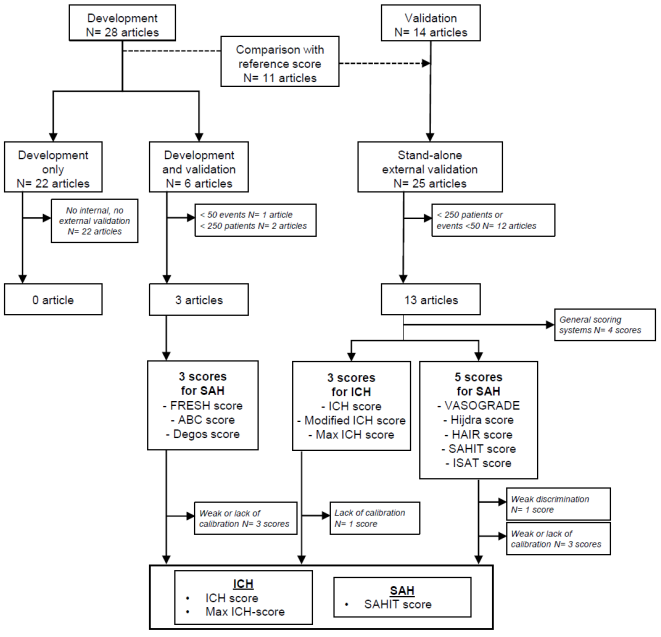

Figure 3 Retained prognostic scores. ICH, intracerebral haemorrhage; SAH, subarachnoid haemorrhages; SAHIT, SubArachnoid Haemorrhage International Trialists.

SAH, 14 (87%) reported a functional outcome, while they represented 6 (46%) of the 13 ICH articles. The primary statistical analysis used to develop the scoring system was logistic regression. Other analyses were linear models, Cox models or less well-known statistical methods such as decision tree analysis, Bayesian networks and artificial neural networks. One article did not specify the type of modelling used (see online supplemental table S3) for corresponding references). Predictor selection strategy, which describes the initial pool of variables and the analysed variables, was rarely mentioned.

Among the 28 included articles, 22 articles developed their tool without validation, that is, they reported the apparent prognostic capacities on the training sample only. They were thus not retained. However, few of these studies were well conducted, with a large cohort and long-term outcome and would deserve validations.[24–27] Among the 28 included articles, six articles presented a development with internal validation (two using bootstrapping, three cross-validations, one temporal validation). One also reported additional external validation. Online supplemental table S5 lists the methods used to quantify prognostic performances. The authors seldom presented global performances. All reported the discrimination with the AUC of the ROC curve, while calibration measures were not systematic. Of the six studies that developed and validated models, two included fewer than 250 patients and one had less than 50 events. We did not retain them due to this insufficient sample size (figure 3).

Finally, three articles proposed a prognostic tool developed and validated based on recommendations: the FRESH score for SAH (excluding rupture of arteriovenous malformation),[11] the ABC score for patients with aneurysmal SAH[28] and the score by Degos *et al* for elderly patients with aneurysmal SAH.[29] Table 1 summarises the collected information regarding source population, development approach, validation details and prognostic performances of these three retained scores.

External validation studies

Fourteen articles aimed to externally validate one or more existing models, most of which were not initially developed with severe injuries managed in ICUs. Eleven out of the 28 articles that developed a tool also compared their score to one or more existing models. Finally, 25 articles presented a stand-alone external validation. Online supplemental table S4 provides complete standardised form and references. Online supplemental table S5 lists the methods used to report prognostic performances. Most reported the AUC of the ROC curve; 15 articles had at least one calibration measurement. The authors rarely compared external validation cohorts to the population of the original article. One study proposed recalibration to predict another outcome than the development study.[30]

Of the 25 studies that externally validated models, 12 included fewer than 250 patients or less than 50 events (figure 3). There were four externally validated general scoring systems. The APACHE II, the SIRS summary score, the SOFA score and the SAPS II showed encouraging performance values when predicting short-term mortality. Because they did not include specific predictors of brain injuries, their use in clinical practice to predict functional or long-term outcomes is not appropriate (figure 3). Injury-specific predictors could extend these scoring systems to improve their predictive capacities and clinical utilities.

There were eight injury-specific externally validated scores. In the ICH population, we retained three externally validated scores: the ICH score,[25 31–33] the modified ICH score (MICH)[25 34] and the max ICH score.[25 33] For the SAH, we retained five tools. Two tools were bivariate, including Glasgow Coma Scale or World Federation of NeuroSurgeons (WFNS) scale associated with CT features: a three-coloured grading system termed the VASOGRADE[35 36] and the Hijdra score for aneurysmal SAH.[37 38] Three tools were multivariate models: the HAIR score[11 36 39 40] and the SubArachnoid Haemorrhage International Trialists (SAHIT) score for SAH,[41 42] and the international subarachnoid aneurysm trial score for aneurysmal SAH.[43 44] Tables 2 and 3 summarise, for ICH and SAH, respectively, the collected information regarding the source population, development approach, validation details and prognostic performances of these eight scores.

Retained prognostic scores from included studies

Finally, for each included study (Development and validation or Stand-alone external validation), we reviewed the levels of prognostic performances for the final selection of multivariate prognostic scores that can be easily applicable for practical use. Among the prognostic tools for

Table 1 Extract of the complete standardised form for the retained developed and validated tools (SAH only)

Author, date, score name	Population	Population characteristics	Sample size, nb of events	Outcomes	Time of prediction, predictors	Discrimination	Calibration	Global performance	Strengths	Limitations
Witsch *et al* 2016 FRESH score[11]	Prospective US monocentric cohort (SHOP). SAH admitted in ICU from 1996 to 2014.	▶ Age: mean 55y. (SD 14) ▶ Female: 38%	n=1526 Unfav. outcome: 79%	Unfav. outcome (mRS 4–6) at 12 m	48 hours after admission. FRESH: Hunt & Hess, APACHE w/o GCS, age, aneurysmal rebleed	AUC: 89.8 (88.1–91.6)	No	▶ N-R^2: 0.50 ▶ C/S-R^2: 0.35	▶ Considering QOL outcome ▶ Good discrimination	▶ Lack of calibration ▶ Use of linear regression to model ordinal mRS
			n=699 Poor cognition: 13%	TICS (cognitive status) at 12 m	48 hours after admission. FRESH-cog: FRESH +education	AUC: 79.7 (75.2–84.2)	No	No		
			n=401 Poor QOL: 11%	SIP (QOL-phys) at 12 m	48 hours after admission. FRESH-quol: FRESH-cog +premorbid disabilities	AUC: 78.2 (71.3–85.2)	No	No		
Degos *et al* 2012 ABC-score[28]	Prospective French monocentric cohort. SAH admitted in NICU from 2003 to 2009.	▶ Age: mean 50y. (SD 13) ▶ Female: 64%	Training n=368 Validation n=158 Mortality: Training 17% Validation 18%	▶ Mortality at 12 m ▶ Independent function (mRS 0–3) at 12 m ▶ Full recovery ▶ (mRS 0–1) at 12 m	On admission. Troponin-I, S100B, GCS	AUC: ▶ Mortality 0.76 (0.67–0.85) ▶ mRS 0–3 0.76 (0.67–0.86) ▶ mRS 0–1 0.83 (0.79–0.88)	HL-GOF: NA	No	▶ Two different cut-offs for mRS ▶ High discrimination	▶ Use of HL-GOF test
Degos *et al* 2012[29]	Prospective French monocentric cohort. SAH admitted in NICU from 2002 to 2010.	▶ Age: <60 y.: 708 60–70 y.: 138 ≥70 y.: 87 ▶ Female: 62%	n=933 Unfav. outcome: 19%	Unfavourable outcome (mRS 4–6) at 12 m	Unknown time of prediction. ▶ Without interaction: IH on admission, severe IH, isch. vasospasm, rebleeding, endovascular complication, surgery complication, Fisher III-V, hydrocephalus, age >60y With interaction: above + hydrocephalus * age >60y	AUC: ▶ Without interaction: 0.84 (0.82–0.88) ▶ With interaction: 0.85 (0.82–0.88)	HL-GOF: ▶ Without interaction: p=0.18 ▶ With interaction: p=0.22	No	▶ High discrimination	▶ Use of HL-GOF test ▶ Unknown time of collection of predictors (endovascular or surgery complications, hydrocephalus)

AP, arterial pressure; AUC, area under the receiving operative curve; AVM, arterio-vascular malformation; compl., complication; GCS, Glasgow Coma Scale; HL-GOF, Hosmer-Lemeshow-Goodness-Of-Fit test; HR, heart rate; ICH, intracerebral haemorrhage; ICU, intensive care unit; IH, intracranial hypertension; isch, ischaemic; m, months; mRS, modified Rankin-Scale; NA, not available; nb, number; NICU, NeuroICU; phys, physical; QOL, quality-of-life; SAH, subarachnoid haemorrhage; TBI, traumatic brain injury; Unfav., unfavourable; y, years.

Table 2 Extract of the complete standardised form for the retained externally validated tools for ICH

Initial paper (author, date), score name	Initial study design	Initial outcomes	Initial predictors	Paper (external validations), date	Design, sample size	Outcomes, nb of events	Discrimination	Calibration	Global performance	Strengths	Limitations
Hemphill *et al* 2001[31] ICH score	American bicentric cohort	Mortality at 1 m	Age, GCS, ICH volume, IVH, Infra-tentorial origin	Rodriguez-Fernandez *et al* 2018 [32]	Prospective multicentre cohort, Spain, 2009–2012 n=336	Mortality at 1 m (52%)	AUC 0.74 (0.69–0.79)	▶ HL-GOF p<0.001 ▶ GiViTI belt p<0.001	No	▶ Good discrimination	▶ Bad calibration ▶ Short-term outcome
				Schmidt *et al* 2018[33]	Prospective US monocentric cohort, 2010–2017 n=372	▶ Mortality at 3 m (41%) ▶ Unfav. outcome (mRS 4–6) at 3 m (63%)	AUC ▶ 3 m mortality 0.83 (0.79–0.88) ▶ 3 m unfav. outcome 0.85 (0.81–0.89)	▶ Histogram pred. vs obs. ▶ HL-GOF p>0.3	Likelihood ratio χ^2 p<0.001	▶ Good discrimination ▶ Functional outcome	▶ Short-term outcome
				Sembill *et al* 2017[25]	Prospective monocentric German cohort, 2007–2011 n=471	▶ Mortality at (3 and) 12 m (30.1%) ▶ Unfav. outcome (mRS 4–6) at (3 and) 12 m (45.4%)	AUC ▶ 12 m mortality 0.69 (0.64–0.74) ▶ 12 m unfav. outcome 0.72 (0.67–0.76)	Histogram pred. vs obs.	No	▶ Long-term outcome ▶ Functional outcome	▶ Weaker discrimination
Cho *et al* 2008[34] MICH	Taiwanese monocentric cohort	▶ Mortality at 6 m ▶ Unfav. outcome (GOS 4–5) at 12 m ▶ Barthel index (≥55) at 12 m	GCS, ICH volume, IVH or hydrocephalus	Sembill *et al* 2017[25]	Prospective monocentric German cohort, 2007–2011 n=471	▶ Mortality at (3 and) 12 m (30.1%) ▶ Unfav. outcome (mRS 4–6) at (3 and) 12 m (45.4%)	AUC ▶ 12 m mortality: 0.6 ▶ 12 m unfav. outcome: 0.69	No	No	Long-term functional outcome	▶ Weak discrimination ▶ No calibration
Sembill *et al* 2017[25] Max ICH score	Germanic monocentric cohort	▶ Mortality at 3 and 12 m ▶ Unfav. outcome (mRS 4–6) at 3 and 12 m	Lobar ICH volume, non-lobar ICH volume, age, NIHSS, IVH, oral anticoag	Schmidt *et al* 2018[33]	Prospective US monocentric cohort, 2010–2017 n=372	▶ Mortality at 3 m (41%) ▶ Unfav. outcome (mRS 4–6) at 3 m (63%)	AUC ▶ 3 m mortality: 0.82 (0.78–0.86) ▶ 3 m unfav. outcome: 0.88 (0.85–0.92)	▶ Histogram pred. vs obs. ▶ HL-GOF p>0.3	Likelihood ratio χ^2 p<0.001	▶ Functional outcome ▶ Good discrimination and calibration	Short-term outcome for external validation

anticoag, anticoagulant therapy; AUC, Area Under the Receiving Operative Curve; GCS, Glasgow Coma Scale; GOS, Glasgow Outcome Scale; HL-GOF, Hosmer-Lemeshow-Goodness-Of-Fit test; ICH, Intracranial Haemorrhage; IVH, intraventricular haemorrhage; m, months; MICH, modified ICH; mRS, modified Rankin-Scale; NIHSS, National Institutes of Health Stroke Scale; obs, observed; pred, predicted; unfav, unfavourable.

Table 3 Extract of the complete standardised form for the retained externally validated tools for SAH

Initial paper (author, date), score name	Initial study design	Initial outcomes	Initial predictors	Paper (external validations), date	Design, sample size	Outcomes, nb of events	Discrimination	Calibration	Global performance	Strengths	Limitations
De Oliveira Manoel *et al* 2015[35] VASOGRADE	3 SAHIT trials (CONSCIOUS-1, EPO trial, statin trial)+1 Canadian centre	DCI	WFNS, modified Fisher Scale	Dengler *et al* 2017[36]	German hosp. registry, monocentric, 2009–2015, n=423	▶ Unfav. outcome (mRS 3–6) at 12 m (53.1%) ▶ Unfav. outcome (mRS 4–6) at 12 m	AUC ▶ mRS 3–6: 0.711 ▶ mRS 4–6: 0.709	Histogram pred. vs obs.	No	▶ Good discrimination ▶ Choice of a functional outcome for EV	▶ Weak calibration for 2/3 of patients due to a three-level ordinal scale with two predictors
Hijdra *et al* 1988[37] Hijdra score	NA	▶ Unfav. outcome (GOS 1 to 2) at 1 m ▶ DCI ▶ Rebleeding	SAH volume, GCS	Claassen *et al* 2004[38]	Prospective US monocentric cohort, 1996–2002, n=413	Unfav. Outcome (mRS 4–6) at 3 m (40.4%)	AUC 0.67 (0.61–0.73)	No	No	▶ Weaker discrimination	▶ No calibration
Risselada *et al* 2010[43] ISAT	ISAT European cohorts	Mortality at 2 m	Age, lumen size, Fisher grade, WFNS	Dijkland *et al* 2016[44]	Dutch hosp. registry, Monocentric, 2007–2011, n=307	Mortality at 2 m (30.6%)	AUC ▶ WFNS at ttt: 0.89 ▶ WFNS on adm: 0.82	Calibration curve ▶ Intercept at trt: 2.248 on adm: 1.502 ▶ Slope at trt: 1.417on adm: 1.959	No	▶ Good discrimination ▶ Calibration curve	▶ Low calibration for high-risk SAH ▶ Choice of the outcome (short-term mortality)
Jaja *et al* 2018[41] SAHIT	SAHIT (nine international trials and registries)	▶ Mortality at 3 m ▶ Unfav. outcome (GOS 1 to 3) at 3 m	▶ Core: age, premorbid history of HT, WFNS ▶ NI: core +CT vol of SAH, aneurysm size, aneurysm location ▶ Full: NI +ttt modality	Mascitelli *et al* 2018[42]	US trial cohort, Monocentric, 2003–2007, n=338	▶ Unfav. outcome (mRS 3–6) at 6 m (29.6%)Mortality at 6 m (10.1%)	AUC ▶ Unfav. outcome: core: 72.8 (66.8–78.9) NI: 73.2 (67.1–79.2) Full: 73.4 (67.5–79.4) ▶ Mortality: core: 72.1 (62.1–82.2) NI: 73.9 (64.4–83.5) Full: 74.4 (65.1–83.8)	Calibration curve intercept/ slope	▶ R^2 ▶ Brier score ▶ Brier scale	▶ Extended choice of predictors ▶ Good discrimination ▶ Good calibration curves ▶ Reporting of global performances	
Lee *et al* 2014[39] HAIR	American monocentric cohort	In-hosp. mortality	Hunt & Hess, age, IVH, rebleeding within 24 hours	Witsch *et al* 2016[11]	▶ Cohort SHOP: Prospective US monocentric cohort, 1996–2014,n=1526Cohort CONSCIOUS –1: multicentric Israel, Europe, North America, n=413	Unfav. outcome (mRS 4–6) at 12 m (79% SHOP) (NA CONSCIOUS-1)	AUC ▶ SHOP: 88.3 (86.4–90.2) ▶ CONSCIOUS-1: 71.8 (66.0–77.5)	No	▶ SHOP: N-R^2: 0.45 C/S-R^2: 0.32 ▶ CONSCIOUS-1: N-R^2: 0.17 C/S-R^2: 0.11	▶ Long-term functional outcome ▶ Good discrimination	▶ Weaker perf. on CONSCIOUS-1 ▶ No calibration
				Dengler *et al*. 2017[36]	Hosp. registry Germany, monocentric 2009–2015, n=423	▶ Unfav. outcome (mRS 3–6) at 12 m (53.1%)Unfav. outcome (mRS 4–6) at 12 months	AUC ▶ mRS 3–6: 0.739 ▶ mRS 4–6: 0.737	Histogram pred. vs obs.	No	▶ Long-term functional outcome ▶ Good discrimination	▶ Weak calibration for high-risk SAH
				Abulhasan *et al* 2017[40]	Canadian retrospective monocentric cohort, 2010–2016, n=434	Mortality at discharge (14.1%)	AUC: 0.89	Calibration curve Intercept: −0.05 Slope: 0.77	No	▶ Excellent discrimination ▶ Calibration curve	Choice of outcome (short-term mortality)

adm, admission; AUC, Area Under the Receiving Operative Curve; C/S-R^2, Cox/Snell-R^2; DCI, Delayed Cerebral Ischaemia; GCS, Glasgow Coma Scale; GOS, Glasgow Outcome Scale; Hosp, hospital; HT, HyperTension; ISAT, international subarachnoid aneurysm trial; IVH, intraventricular haemorrhage; m, months; mRS, modified Rankin-Scale; NA, Not available; NI, NeuroImaging; N-R^2, Nagelkerke-R^2; obs, observed; pred, predicted; SAH, SubArachnoid Haemorrhage; SAHIT, SubArachnoid Haemorrhage International Trialists; trt, treatment; unfav., unfavourable; vol, volume; WFNS, World Federation of NeuroSurgeons.

ICH, we did not retain the MICH score because of the lack of reporting calibration that did not guarantee agreement between predictions and observed outcomes. We thus highlighted two scores (figure 3). The ICH score[31] was externally validated in three large ICU cohorts, predicting 1-month, 3-month and 12-month mortality or functional outcome (mRS 4–6).[25 32 33] The max ICH score[25] predicted 3-month and 12-month mortality and functional outcome (mRS 4–6), based on CT predictors (lobar and non-lobar ICH volume, age, National Institutes of Health Stroke Scale, presence of intraventricular haemorrhage and anticoagulant therapy). This showed good performances in a large external ICU cohort.[33] Table 2 presents the original publication, the external validation studies and corresponding performances (discrimination and calibration).

Among the retained SAH tools, the level of clinical utility and prognostic capacities was debatable. Tables 1 and 3 detail the strengths and limitations of each of these scores. The vast majority of tools presented high discrimination. We did not retain the Hijdra score[37] because of weak discrimination or absence of calibration. Additionally, the VASOGRADE,[35] the FRESH score,[11] the ABC score[28] and the Degos score for the elderly[29] lacked reporting calibration or used the Hosmer-Lemeshow goodness-of-fit test. The ISAT score and HAIR score, which had a low calibration for high-risk SAH, would probably benefit from recalibration or updating.[39 43] We thus only retained the SAHIT score[41] (figure 3). In a single external validation,[42] it predicted either an unfavourable outcome (mRS 3–6) or mortality at 6 months, based on clinical predictors (age, history of hypertension and WFNS preoperative neurological grade) and CT (Fisher grade, aneurysm size and location). It revealed good discrimination and calibration.

DISCUSSION

While studies labelled as 'prognostic' abound in the literature on intracranial haemorrhage, our systematic review dedicated explicitly to critical patients revealed a lack of methodological robustness. Of the 85 read articles, we identified six articles that developed a prognostic tool supported by a validation study and 25 external validation studies. After critical appraisal of the articles, we retained, for the ICH population, the ICH score,[31] which has better performances for the shorter outcome, and the max ICH score.[25] For the SAH population, we retained the SAHIT score for its high methodological quality.[42]

The ICH score,[31] developed in 2001, has benefited from multiple external validations in many different populations. The American Heart Association guidelines[45] recommend its reporting. In external validations with severe ICH, its performances could be better, particularly for longer term and functional outcomes.[25 32 33] It would be interesting to consider updating or recalibrating this tool. The max ICH score,[25] developed in 2017, showed good calibration and discrimination on only one external cohort, with satisfying calibration and better performances than the ICH score on the same sample.[33] It would benefit from further validations in other large and recent cohorts. The SAHIT score, developed in 2018, predicted unfavourable outcome or mortality at 3 months in a low to severe SAH population.[41] The single external validation in an ICU cohort revealed good prognostic performances that further studies have yet to be confirmed.[42]

In our systematic review, the authors rarely highlighted the clinical objective, which leads us to believe that clinical purposes did not drive most score elaborations. Functional outcomes in the modern setting of critical care make more sense than mortality outcomes for patients who are more likely to survive but face disabilities.[46] The ordinal functional outcomes scales are almost systematically dichotomised (GOS 1–3 vs 4–5, mRS 4–6 or 3–6 vs 0–3 or 0–4). These thresholds, though never justified, should depend on the clinical objective. If the score's purpose is to support clinicians in making ethically challenging decisions, such as withdrawal of care, it is not reasonable to place severe disabilities, vegetative state and death on the same unfavourable side. Besides, a prognostic tool on its own, as rigorous as it may be, is hardly capable of integrating the strong human dimension of such a complex decision. Multidisciplinary clinical teams should rely on a combination of considerations, which include multivariable scoring systems. If the clinical objective is instead to inform patients and their relatives of the evolution prospects, the condition they consider to be favourable should be determined by themselves and ideally over the very long-term.[47 48] In our systematic review, the longest prediction horizon was 12 months, that is, before stabilisation of functional recovery and the ability to adapt to such a consolidated statement.[49] Moreover, patient perception could weigh the different levels of functional disabilities.[50 51] Indeed, survivors have a wide range of lifelong consequences such as neuropsychological difficulties, memory problems, fatigue and physical complaints, that is, dimensions not explored with functional outcome scales.[51 52] As these symptoms are not always apparent, only validated patient (or caregivers) reported questionnaires can reflect the subjective perception of their quality of life.[51 53] In our systematic review, the only article mentioning quality of life concerns the FRESH score.[11] Even though some methodological choices are questionable in this study, we think that it deserves attention because it surpasses the functional outcomes by integrating the quality of life as an objective of prediction.

In our systematic review, we identified several methodological pitfalls. A large proportion of eligible studies are wrongly labelled 'prognostic models'. Some authors did not report prognostic performances, sometimes because they wrongly interpreted the odds ratio as a prognostic ability. These mistakes revealed considerable confusion in the literature between the notions of correlation and prediction.[54] Some development studies only reported apparent prognostic performances. This lack of internal or external validation led to overestimating the

performances of the prognostic tools.[21] Several studies based on small sample size or a small number of events resulted in the risk of overfitting or low credibility in terms of prognostic performances.[55] These studies would benefit from external validations with recent and large cohorts. There was heterogeneity in the prognostic performances' reports: discrimination was systematic, only about half of the retained studies assessed calibration and 10% global performance. Calibration curves, rarely reported, allow future external validation to assess the eventual need for recalibration or updating, to adapt it to the population of interest. The popular Hosmer-Lemeshow Goodness-of-fit test is known to perform poorly, making its use regrettable.[56] We discarded several studies due to the absence of variables selection. The included studies rarely specified the predictor selection strategy, which describes the initial pool of variables and the analysed variables. This precision allows the reader to assess the risk of overfitting. The prediction time was sometimes unknown, making the score challenging to apply. Authors should clearly state this information to inform the user of when to calculate the prediction. Authors who studied long-term outcomes always chose to use logistic regression by excluding patients lost to follow-up when times-to-event methods would have been more appropriate in the presence of such censoring. Finally, this resulted in a very low number of prognostic tools that seemed methodologically correct and presenting a reasonable prognostic performance level. However, weaker validation results do not mean that the model is incorrect. If scores' development approaches were optimal, relevant predictors could be recalibrated and combined with new data to validate a strong tool.[57 58]

A consortium of experts published the TRIPOD statement in 2015, clearly setting out how to report prognostic information.[17] Of the 85 full texts screened in our review, 35 (41%) were printed after the TRIPOD publication in 2015. Of these, we finally retained 19 (54%) articles published after the TRIPOD publication, whereas we retained only 11 (22%) from the 50 articles published before the TRIPOD publication. Similar to Zamanipoor Najafabadi *et al*, we noticed a trend towards quality improvement, reinforced with the necessary ongoing validation of existing scores.[59] Our systematic review revealed that some robust published scores, outlined in reviews focusing on non-severe to severe intracranial haemorrhages[10–13] (eg, FUNC score,[60] Essen ICH score[61] and ICHOP score[62]), have not yet been validated in the ICU population. To use them reliably in such settings, they should be externally validated with critical patients. We also did not find tools dedicated to severe specific populations (such as haemorrhages secondary to malformation or patients with coagulation disorders). External validations would be interesting for these populations (eg, patients under anticoagulant[63] or arteriovenous malformation[64]). Another option would be to extend existing scores with these risk predictors, such as the Max ICH-score, which includes the variable 'presence of oral anticoagulant'.[25] With the rapid evolution of therapeutic advances in neurocritical care, the ongoing prognostic studies should focus on temporal validation and updating/recalibrating existing good scores to ensure their performance validity.[55] It is also possible to extend this by incorporating additional modern variables.

This review has several limitations. First, we aimed to include tools dedicated to ICH or SAH managed in the ICU. Because of the lack of severity classification for these pathologies, and heterogeneity of patients admitted to ICUs, we defined our proper severity criteria, which is debatable. Second, only one assessor conducted the study screening on title/abstract. This may have resulted in some missing eligible studies. Third, we did not use a formal tool to study the risk of bias such as the recent Prediction model Risk Of Bias ASsessment Tool (PROBAST) based on the TRIPOD.[65 66] Following the TRIPOD recommendations, we built our own standardised form collecting similar information than the PROBAST items. Fourth, due to the heterogeneity in the included models, we could not to perform a meta-analysis. Finally, as with any systematic review, our work underwent publication bias issue. Similar to randomised clinical trials, we cannot exclude that unpublished studies may have negative results or size effects different from published studies.[67] One consequence could be, for instance, the underrepresentation of external validation studies with non-confirmatory prognostic performances.

CONCLUSIONS

Our review identified several methodological pitfalls and incomplete reporting in prognostic articles on intracranial haemorrhages managed in ICU. Among the many published scores for ICH and SAH, some deserve further attention. Rather than developing new scores, future authors should focus on externally validating and updating well-developed existing scores with large and recent cohorts, relying on methodological syntheses such as the TRIPOD statement.[17 57 68] We have chosen to emphasise the ICH score, the max ICH score and the SAHIT scores for their superior prognostic performances. Nevertheless, they need ongoing validations, recalibrations and impact studies to improve them. The use of 'patient-centred' outcomes that have yet to be defined could also enhance the tools in the delicate, medical and ethical setting of critical care. Beyond all methodological issues, patient-centred clinical finality should guide prognostic tools to be convincing.

Author affiliations
[1]UMR 1246 Methods in Patients-Centered Outcomes and Health Research, INSERM, Nantes, France
[2]Biostatistician, University Hospital Centre Nantes, Nantes, Pays de la Loire, France
[3]Medical Intensive Care, Angers University Hospital, Nantes, France
[4]Biostatistics and Methodology Unit, University Hospital Centre Nantes, Nantes, Pays de la Loire, France
[5]Intensive Care Unit, Regional University Hospital Centre Tours, Tours, Centre, France

[6]Anaesthesia and Intensive Care Unit, Hôpital Laennec, Saint-Herblain, University Hospital of Nantes, France, Université de Nantes, CHU Nantes, Saint-Herblain, France
[7]Anesthesia and Critical Care Department, University Hospital Centre Poitiers, Poitiers, France

Twitter Jeanne Simon-Pimmel @PimmelSimon

Contributors JS-P: conceptualisation and design of the review, data acquisition, statistical analysis, writing: original draft, review and editing. YF: conceptualisation, data acquisition, providing statistical expertise, writing: original draft, review and editing. FF: data acquisition, writing: review and editing. ML, LB-C, RC and DF: data acquisition, providing clinical expertise, writing: review and editing. ED: project management, conceptualisation, methodology, data acquisition, writing: original draft, review and editing.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests YF reports personal fees for statistical training from Sanofi and Biogen, outside the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Jeanne Simon-Pimmel http://orcid.org/0000-0003-0699-5793
Etienne Dantan http://orcid.org/0000-0001-7137-5051

REFERENCES

1 van Gijn J, Kerr RS, Rinkel GJE. Subarachnoid haemorrhage. *Lancet* 2007;369:306–18.
2 Balami JS, Buchan AM. Complications of intracerebral haemorrhage. *Lancet Neurol* 2012;11:101–18.
3 van Asch CJ, Luitse MJ, Rinkel GJ, *et al*. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol* 2010;9:167–76.
4 Rinkel GJE, Algra A. Long-Term outcomes of patients with aneurysmal subarachnoid haemorrhage. *Lancet Neurol* 2011;10:349–56.
5 Katrak PH, Black D, Peeva V. Do stroke patients with intracerebral hemorrhage have a better functional outcome than patients with cerebral infarction? *Pm R* 2009;1:427–33.
6 Hamel MB, Goldman L, Teno J. Identification of comatose patients at high risk for death or severe disability. *JAMA* 1995;273:1842–8.
7 Pratt AK, Chang JJ, Sederstrom NO. A fate worse than death: prognostication of devastating brain injury. *Crit Care Med* 2019;47:591–8.
8 Gillett GR, Honeybul S, Ho KM, *et al*. Neurotrauma and the rub: where tragedy meets ethics and science. *J Med Ethics* 2010;36:727–30.
9 Jain A, Jain M, Bellolio MF, *et al*. Is early DNR a self-fulfilling prophecy for patients with spontaneous intracerebral hemorrhage? *Neurocrit Care* 2013;19:342–6.
10 Jaja BNR, Cusimano MD, Etminan N, *et al*. Clinical prediction models for aneurysmal subarachnoid hemorrhage: a systematic review. *Neurocrit Care* 2013;18:143–53.
11 Witsch J, Frey H-P, Patel S, *et al*. Prognostication of long-term outcomes after subarachnoid hemorrhage: the fresh score. *Ann Neurol* 2016;80:46–58.
12 Sreekrishnan A, Dearborn JL, Greer DM, *et al*. Intracerebral hemorrhage location and functional outcomes of patients: a systematic literature review and meta-analysis. *Neurocrit Care* 2016;25:384–91.
13 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.
14 Gregório T, Pipa S, Cavaleiro P, *et al*. Prognostic models for intracerebral hemorrhage: systematic review and meta-analysis. *BMC Med Res Methodol* 2018;18:145.
15 Hemingway H, Croft P, Perel P, *et al*. Prognosis research strategy (progress) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.
16 Steyerberg EW, Moons KGM, van der Windt DA, *et al*. Prognosis research strategy (progress) 3: prognostic model research. *PLoS Med* 2013;10:e1001381.
17 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
18 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264–9.
19 Moons KGM, de Groot JAH, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.
20 Moons KGM, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
21 Steyerberg EW, Bleeker SE, Moll HA, *et al*. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol* 2003;56:441–7.
22 Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;35:214–26.
23 Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Stat Methods Med Res* 2017;26:796–808.
24 Fallenius M, Skrifvars MB, Reinikainen M, *et al*. Spontaneous intracerebral hemorrhage. *Stroke* 2019;50:2336–43.
25 Sembill JA, Gerner ST, Volbers B, *et al*. Severity assessment in maximally treated ICH patients: the max-ICH score. *Neurology* 2017;89:423–31.
26 Weimer JM, Nowacki AS, Frontera JA. Withdrawal of life-sustaining therapy in patients with intracranial hemorrhage: self-fulfilling prophecy or accurate prediction of outcome? *Crit Care Med* 2016;44:1161–72.
27 Zheng K, Zhong M, Zhao B, *et al*. Poor-Grade aneurysmal subarachnoid hemorrhage: risk factors affecting clinical outcomes in intracranial aneurysm patients in a multi-center study. *Front Neurol* 2019;10:123.
28 Degos V, Apfel CC, Sanchez P, *et al*. An admission bioclinical score to predict 1-year outcomes in patients undergoing aneurysm coiling. *Stroke* 2012;43:1253–9.
29 Degos V, Gourraud P-A, Tursis VT, *et al*. Elderly age as a prognostic marker of 1-year poor outcome for subarachnoid hemorrhage patients through its interaction with admission hydrocephalus. *Anesthesiology* 2012;117:1289–99.
30 Fallenius M, Skrifvars MB, Reinikainen M, *et al*. Common intensive care scoring systems do not outperform age and Glasgow coma scale score in predicting mid-term mortality in patients with spontaneous intracerebral hemorrhage treated in the intensive care unit. *Scand J Trauma Resusc Emerg Med* 2017;25:102.
31 Hemphill JC, Bonovich DC, Besmertis L, *et al*. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke* 2001;32:891–7.
32 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.
33 Schmidt FA, Liotta EM, Prabhakaran S, *et al*. Assessment and comparison of the max-ICH score and ICH score by external validation. *Neurology* 2018;91:e939–46.
34 Cho D-Y, Chen C-C, Lee W-Y, *et al*. A new Modified Intracerebral Hemorrhage score for treatment decisions in basal ganglia hemorrhage--a randomized trial. *Crit Care Med* 2008;36:2151–6.

35 de Oliveira Manoel AL, Jaja BN, Germans MR, *et al*. The VASOGRADE: a simple grading scale for prediction of delayed cerebral ischemia after subarachnoid hemorrhage. *Stroke* 2015;46:1826–31.
36 Dengler NF, Sommerfeld J, Diesing D, *et al*. Prediction of cerebral infarction and patient outcome in aneurysmal subarachnoid hemorrhage: comparison of new and established radiographic, clinical and combined scores. *Eur J Neurol* 2018;25:111–9.
37 Hijdra A, van Gijn J, Nagelkerke NJ, *et al*. Prediction of delayed cerebral ischemia, rebleeding, and outcome after aneurysmal subarachnoid hemorrhage. *Stroke* 1988;19:1250–6.
38 Claassen J, Vu A, Kreiter KT, *et al*. Effect of acute physiologic derangements on outcome after subarachnoid hemorrhage. *Crit Care Med* 2004;32:832–8.
39 Lee VH, Ouyang B, John S, *et al*. Risk stratification for the in-hospital mortality in subarachnoid hemorrhage: the hair score. *Neurocrit Care* 2014;21:14–19.
40 Abulhasan YB, Alabdulraheem N, Simoneau G, *et al*. Mortality after spontaneous subarachnoid hemorrhage: causality and validation of a prediction model. *World Neurosurg* 2018;112:e799–811.
41 Jaja BNR, Saposnik G, Lingsma HF, *et al*. Development and validation of outcome prediction models for aneurysmal subarachnoid haemorrhage: the SAHIT multinational cohort study. *BMJ* 2018;360:j5745.
42 Mascitelli JR, Cole T, Yoon S, *et al*. External validation of the subarachnoid hemorrhage international Trialists (SAHIT) predictive model using the Barrow ruptured aneurysm trial (Brat) cohort. *Neurosurgery* 2020;86:101–6.
43 Risselada R, Lingsma HF, Bauer-Mehren A, *et al*. Prediction of 60 day case-fatality after aneurysmal subarachnoid haemorrhage: results from the International subarachnoid aneurysm trial (ISAT). *Eur J Epidemiol* 2010;25:261–6.
44 Dijkland SA, Roozenbeek B, Brouwer PA, *et al*. Prediction of 60-Day case fatality after aneurysmal subarachnoid hemorrhage: external validation of a prediction model. *Crit Care Med* 2016;44:1523–9.
45 Hemphill JC, Greenberg SM, Anderson CS, *et al*. Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the American heart Association/ American stroke association. *Stroke* 2015;46:2032–60.
46 Peigne V, Chaize M, Falissard B, *et al*. Important questions asked by family members of intensive care unit patients. *Crit Care Med* 2011;39:1365–71.
47 Saigle V, Asad S, Presseau J, *et al*. Do patient-reported outcome measures for SAH include patient, family, and caregiver priorities? A scoping review. *Neurology* 2019;92:281–95.
48 Gaudry S, Messika J, Ricard J-D, *et al*. Patient-important outcomes in randomized controlled trials in critically ill patients: a systematic review. *Ann Intensive Care* 2017;7:28.
49 Ashley JG, Ashley MJ, Masel BE, *et al*. The influence of post-acute rehabilitation length of stay on traumatic brain injury outcome: a retrospective exploratory study. *Brain Inj* 2018;32:600–7.
50 Chaisinanunkul N, Adeoye O, Lewis RJ, *et al*. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a Utility-Weighted modified Rankin scale. *Stroke* 2015;46:2238–43.
51 Nobels-Janssen E, van der Wees PJ, Verhagen WIM, *et al*. Patient-Reported outcome measures in subarachnoid hemorrhage: a systematic review. *Neurology* 2019;92:1096–112.
52 Tagliaferri F, Compagnone C, Korsic M, *et al*. A systematic review of brain injury epidemiology in Europe. *Acta Neurochir* 2006;148:255–68.
53 Reeves M, Lisabeth L, Williams L, *et al*. Patient-Reported outcome measures (PROMs) for acute stroke: rationale, methods and future directions. *Stroke* 2018;49:1549–56.
54 Pepe MS, Janes H, Longton G, *et al*. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;159:882–90.
55 Steyerberg EW, Uno H, Ioannidis JPA, *et al*. Poor performance of clinical prediction models: the harm of commonly applied methods. *J Clin Epidemiol* 2018;98:133–43.
56 Hosmer DW, Hosmer T, Le Cessie S, *et al*. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965–80.
57 Van Calster B, Van Hoorde K, Vergouwe Y, *et al*. Validation and updating of risk models based on multinomial logistic regression. *Diagn Progn Res* 2017;1:2.
58 Steyerberg EW. Validation in prediction research: the waste by data splitting. *J Clin Epidemiol* 2018;103:131–3.
59 Zamanipoor Najafabadi AH, Ramspek CL, Dekker FW, *et al*. Tripod statement: a preliminary pre-post analysis of reporting and methods of prediction models. *BMJ Open* 2020;10:e041537.
60 Rost NS, Smith EE, Chang Y, *et al*. Prediction of functional outcome in patients with primary intracerebral hemorrhage: the FUNC score. *Stroke* 2008;39:2304–9.
61 Weimar C, Roth M, Willig V, *et al*. Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage. *J Neurol* 2006;253:788–93.
62 Gupta VP, Garton ALA, Sisti JA, *et al*. Prognosticating functional outcome after intracerebral hemorrhage: the ICHOP score. *World Neurosurg* 2017;101:577–83.
63 Katsanos AH, Krogias C, Lioutas V-A, *et al*. The prognostic utility of ICH-score in anticoagulant related intracerebral hemorrhage. *J Neurol Sci* 2020;409:116628.
64 Appelboom G, Hwang BY, Bruce SS, *et al*. Predicting outcome after arteriovenous malformation-associated intracerebral hemorrhage with the original ICH score. *World Neurosurg* 2012;78:646–50.
65 Wolff RF, Moons KGM, Riley RD, *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.
66 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.
67 Rezaeian M. Reducing publication bias in biomedical research: reviewing and registering protocol with a suitable Journal. *J Clin Epidemiol* 2016;69:248–9.
68 Collins GS, de Groot JA, Dutton S, *et al*. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 2014;14:40.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 6
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pages 6-7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 7
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 7
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 7, File S2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 8
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 8
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 8
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 8
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Pages 8-9
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 9
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	NA
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	NA
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	NA
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	NA
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	NA
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	NA
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Pages 7-8
Certainty	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Pages 8-9,

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			CHARMS & TRIPOD statement
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 9, Figures 1 & 3
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Page 9, Figures 1 & 3
Study characteristics	17	Cite each included study and present its characteristics.	Pages 9-13, Tables 1-3 & S3-S4
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Pages 9-13, Tables 1-3 & S3-S4
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Pages 9-13, Tables 1-3 & S3-S4
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	NA
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	NA
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Pages 9-13, Tables 1-3 & S3-S4
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Pages 13-14, Figure 3
DISCUSSION			

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pages 14-15
	23b	Discuss any limitations of the evidence included in the review.	Pages 16-17
	23c	Discuss any limitations of the review processes used.	Pages 17-18
	23d	Discuss implications of the results for practice, policy, and future research.	Page 18
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Not registered
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Not registered
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not registered
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 19
Competing interests	26	Declare any competing interests of review authors.	Page 19
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Template data collection form on demand, Tables 1-3 & S3-S4

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

S2 File: Detailed searched strategy

Medline (via Pubmed)

("Brain Injuries"[MeSH Terms] OR "Brain Injuries"[Title/Abstract] OR "Brain Injury"[Title/Abstract] OR "Injury, Brain"[Title/Abstract] OR "Injuries, Brain"[Title/Abstract] OR "Brain Injuries, Diffuse"[Title/Abstract] OR "Brain Injury, Diffuse"[Title/Abstract] OR "Diffuse Brain Injuries"[Title/Abstract] OR "Diffuse Brain Injury"[Title/Abstract] OR "Injuries, Diffuse Brain"[Title/Abstract] OR "Injury, Diffuse Brain"[Title/Abstract] OR "Brain Injuries, Focal"[Title/Abstract] OR "Brain Injury, Focal"[Title/Abstract] OR "Focal Brain Injury"[Title/Abstract] OR "Injuries, Focal Brain"[Title/Abstract] OR "Injury, Focal Brain"[Title/Abstract] OR "Focal Brain Injuries"[Title/Abstract] OR "Injuries, Acute Brain"[Title/Abstract] OR "Acute Brain injury"[Title/Abstract] OR "Brain Injury, Acute"[Title/Abstract] OR "Injury, Acute Brain"[Title/Abstract] OR "Brain Injuries, Acute"[Title/Abstract] OR "Acute Brain Injuries"[Title/Abstract] OR "Brain Hemorrhage"[Title/Abstract] OR "Brain Hemorrhages"[Title/Abstract] OR "Brain Haemorrhage"[Title/Abstract] OR"Brain Haemorrhages"[Title/Abstract]

OR
"Traumatic Brain Injury"[Title/Abstract] OR "Traumatic Brain Injuries"[Title/Abstract] OR "Brain Injuries, Traumatic"[Title/Abstract] OR "Brain Injury, Traumatic"[Title/Abstract] OR "Injuries, Traumatic Brain"[Title/Abstract] OR "Injury, Traumatic Brain"[Title/Abstract] OR "Injury, Brain, Traumatic"[Title/Abstract] OR "Brain Hemorrhage, Traumatic"[MeSH Terms] OR "Brain Lacerations"[Title/Abstract] OR "Brain Laceration"[Title/Abstract] OR "Laceration, Brain"[Title/Abstract] OR "Lacerations, Brain"[Title/Abstract] OR "Cortical Contusion"[Title/Abstract] OR "Contusion, Cortical"[Title/Abstract] OR "Contusions, Cortical"[Title/Abstract] OR "Cortical Contusions"[Title/Abstract] OR "Brain Contusion"[Title/Abstract] OR "Brain Contusions"[Title/Abstract] OR "Contusion, Brain"[Title/Abstract] OR "Contusions, Brain"[Title/Abstract] OR "Trauma, Brain"[Title/Abstract] OR "Brain Trauma"[Title/Abstract] OR "Brain Traumas"[Title/Abstract] OR "Traumas, Brain"[Title/Abstract] OR "TBI"[Title/Abstract] OR "TBIs"[Title/Abstract]

OR
"Intracranial Arterial Diseases"[MeSH Terms] OR "Intracranial Arterial Diseases"[Title/abstract] OR "Intracranial Arterial Disease"[Title/Abstract] OR "Intracranial Aneurysm"[MeSH Terms] OR "Intracranial Aneurysm"[Title/Abstract] OR "Intracranial Arteriovenous Malformations"[MeSH Terms] OR "Intracranial Arteriovenous Malformations"[Title/Abstract] OR "Intracranial Arteriovenous Malformation"[Title/Abstract] OR "Intracranial Hemorrhages"[MeSH Terms] OR "Intracranial Hemorrhages"[Title/Abstract] OR "Intracranial Haemorrhages"[Title/Abstract] OR "Intracranial Hemorrhage"[Title/Abstract] OR "Intracranial Haemorrhage"[Title/Abstract] OR "Cerebral Hemorrhage"[MeSH Terms] OR "Cerebral Hemorrhage"[Title/Abstract] OR "Cerebral Haemorrhage"[Title/Abstract] OR "Cerebral Hemorrhages"[Title/Abstract] OR "Cerebral Haemorrhages"[Title/Abstract] OR "Subarachnoid Hemorrhage"[MeSH Terms] OR "Subarachnoid Hemorrhage"[Title/Abstract] OR "Subarachnoid Haemorrhage"[Title/Abstract] OR "Subarachnoid Hemorrhages"[Title/Abstract] OR "Subarachnoid Haemorrhages"[Title/Abstract])

AND
("predictive value of tests"[MeSH Terms] OR "prediction"[Title/Abstract] OR "predict"[Title/Abstract] OR "predicts"[Title/Abstract] OR "predictive"[Title/Abstract] OR "predicting"[Title/Abstract] OR "predicted"[Title/Abstract] OR "prognosis"[MeSH Terms] OR "prognosis"[Title/Abstract] OR "prognoses"[Title/Abstract] OR "prognostication"[Title/Abstract] OR "prognosticate"[Title/Abstract] OR "prognosticates"[Title/Abstract] OR "prognostic"[Title/Abstract])

AND
("models, statistical"[MeSH Terms] OR "model"[Title/Abstract] OR "models"[Title/Abstract] OR "modeling"[Title/Abstract] OR "modelling"[Title/Abstract] OR "equation"[Title/Abstract] OR "equations"[Title/Abstract] OR "regression"[Title/Abstract] OR "algorithm"[Title/Abstract] OR "algorithms"[Title/Abstract] OR "score"[Title/Abstract] OR "scores"[Title/Abstract] OR "scoring"[Title/Abstract] OR "nomograms"[MeSH Terms] OR "nomograms"[Title/Abstract] OR "nomogram"[Title/Abstract] OR "rule" [Title/Abstract] OR "rules" [Title/Abstract])

AND

("validation studies as topic"[MeSH Terms] OR "validation"[Title/Abstract] OR "validity"[Title/Abstract] OR "validate"[Title/Abstract] OR "validates"[Title/Abstract] OR "validated"[Title/Abstract] OR "calibration"[MeSH Terms] OR "calibration"[Title/Abstract] OR "calibrated"[Title/Abstract] OR "calibrates"[Title/Abstract] OR "calibrate"[Title/Abstract] OR "discrimination"[Title/Abstract] OR "discriminates"[Title/Abstract] OR "discriminate"[Title/Abstract] OR "discriminated"[Title/Abstract] OR "concordance"[Title/Abstract] OR "C-statistic"[Title/Abstract] OR "C-index"[Title/Abstract] OR "ROC curve"[MeSH Terms] OR "ROC"[Title/Abstract] OR "area under curve"[MeSH Terms] OR "AUC"[Title/Abstract] or "development" [Title])

NOT
("preterm"[Title/Abstract] OR "neonatal"[Title/Abstract] OR "Infant, Premature"[MeSH Terms] OR "Premature"[Title/Abstract] OR "Infant"[MeSH Terms] OR "Infant"[Title/Abstract] OR "Infants"[Title/Abstract] OR "Infant, Newborn"[Mesh Terms] OR "Newborn"[Title/Abstract] OR "Newborns"[Title/Abstract] OR "neonate"[Title/Abstract] OR "neonates"[Title/Abstract])

COCHRANE LIBRARY

#1 brain injuri* or brain h*morrhage* or traumatic brain injur* or brain laceration* or Cortical Contusion* or Brain Contusion* or Brain Trauma* or TBI* or Intracranial Arterial Disease* or Intracranial Aneurysm* or Intracranial Arteriovenous Malformation* or Intracranial H*morrhage* or Cerebral H*morrhage* or Subarachnoid H*morrhage
#2 predict* or prognos*
#3 model* or equation* or regression* or algorithm* or rule* or scor* or nomogram*
#4 validat* or calibrat* or discriminat* or concordance or C-statistic or C-index or ROC or AUC or development
#5 preterm OR neonat* OR infant* OR prematur* OR newborn
#6
#1 and #2 and #3 and #4 not #5

WEB OF SCIENCE

TS=(brain injur* OR brain hemorrhage* OR brain haemorrhage* OR traumatic brain injur* OR brain laceration* OR Cortical Contusion* OR Brain Contusion* OR Brain Trauma* OR TBI*
OR
Intracranial Arterial Disease* OR Intracranial Aneurysm OR Intracranial Arteriovenous Malformation* OR Intracranial Hemorrhage* OR Intracranial Haemorrhage* OR Cerebral Hemorrhage* OR Cerebral Haemorrhage* OR Subarachnoid Hemorrhage* OR Subarachnoid Haemorrhage*)
AND TS=(predict* OR prognos*)
AND TS=(model* OR equation* OR regression* OR algorithm* OR rule* OR scor* OR nomogram*)
AND (TS=(valid* OR calibrat* OR discriminat* OR concordance OR C-statistic OR C-index OR ROC OR AUC OR area under curve) OR TI=(development))
NOT TS=(preterm OR neonatal OR infant* OR premature OR newborn OR neonate)

EMBASE

('brain injuries'/exp OR 'brain injuries' OR 'brain injury'/exp OR 'brain injury' OR 'injury, brain'/exp OR 'injury, brain' OR 'injuries, brain' OR 'brain injuries, diffuse'/exp OR 'brain injuries, diffuse' OR 'brain injury, diffuse' OR 'diffuse brain injuries' OR 'diffuse brain injury'/exp OR 'diffuse brain injury' OR 'injuries, diffuse brain' OR 'injury, diffuse brain' OR 'brain injuries, focal' OR 'brain injury, focal' OR 'focal brain injury' OR 'injuries, focal brain' OR 'injury, focal brain' OR 'focal brain injuries' OR 'injuries, acute brain' OR 'acute brain injury'/exp OR 'acute brain injury' OR 'brain injury, acute' OR 'injury, acute brain' OR 'brain injuries, acute' OR 'acute brain injuries' OR 'brain hemorrhage'/exp OR 'brain hemorrhage' OR 'brain hemorrhages' OR 'brain haemorrhage'/exp OR 'brain haemorrhage' OR 'brain haemorrhages' OR 'traumatic brain injury'/exp OR 'traumatic brain injury' OR 'traumatic brain injuries'/exp OR 'traumatic brain injuries' OR 'brain injuries, traumatic'/exp OR 'brain injuries, traumatic' OR 'brain injury, traumatic' OR 'injuries, traumatic brain' OR 'injury, traumatic brain' OR 'injury, brain, traumatic' OR 'brain hemorrhage, traumatic'/exp OR 'brain hemorrhage, traumatic' OR 'brain lacerations' OR 'brain laceration'/exp OR 'brain laceration' OR 'laceration, brain' OR 'lacerations, brain' OR 'cortical contusion' OR 'contusion, cortical' OR 'contusions, cortical' OR 'cortical contusions' OR 'brain contusion'/exp OR 'brain contusion' OR 'brain contusions' OR 'contusion, brain'/exp OR 'contusion, brain' OR 'contusions, brain' OR 'trauma, brain' OR 'brain trauma'/exp OR 'brain trauma' OR 'brain traumas' OR 'traumas, brain' OR 'tbi' OR 'tbis' OR 'intracranial arterial diseases'/exp OR 'intracranial arterial diseases' OR 'intracranial arterial disease' OR 'intracranial aneurysm'/exp OR 'intracranial aneurysm' OR 'intracranial arteriovenous malformations'/exp OR 'intracranial arteriovenous malformations' OR 'intracranial arteriovenous malformation' OR 'intracranial

hemorrhages'/exp OR 'intracranial hemorrhages' OR 'intracranial haemorrhages'/exp OR 'intracranial haemorrhages' OR 'intracranial hemorrhage'/exp OR 'intracranial hemorrhage' OR 'intracranial haemorrhage'/exp OR 'intracranial haemorrhage' OR 'cerebral hemorrhage'/exp OR 'cerebral hemorrhage' OR 'cerebral haemorrhage'/exp OR 'cerebral haemorrhage' OR 'cerebral hemorrhages' OR 'cerebral haemorrhages' OR 'subarachnoid hemorrhage'/exp OR 'subarachnoid hemorrhage' OR 'subarachnoid haemorrhage'/exp OR 'subarachnoid haemorrhage' OR 'subarachnoid hemorrhages' OR 'subarachnoid haemorrhages')

AND
('predictive value of tests'/exp OR 'predictive value of tests' OR 'prediction'/exp OR 'prediction' OR 'predict' OR 'predicts' OR 'predictive' OR 'predicting' OR 'predicted' OR 'prognosis'/exp OR 'prognosis' OR 'prognoses' OR 'prognostication' OR 'prognosticate' OR 'prognosticates' OR 'prognostic')

AND
('models, statistical'/exp OR 'models, statistical' OR 'model'/exp OR 'model' OR 'models'/exp OR 'models' OR 'modeling'/exp OR 'modeling' OR 'modelling'/exp OR 'modelling' OR 'equation'/exp OR 'equation' OR 'equations' OR 'regression'/exp OR 'regression' OR 'algorithm'/exp OR 'algorithm' OR 'algorithms'/exp OR 'algorithms' OR 'score'/exp OR 'score' OR 'scores' OR 'scoring' OR 'nomograms'/exp OR 'nomograms' OR 'nomogram'/exp OR 'nomogram' OR 'rule' OR 'rules') AND ('validation studies as topic'/exp OR 'validation studies as topic' OR 'validation'/exp OR 'validation' OR 'validity'/exp OR 'validity' OR 'validate' OR 'validates' OR 'validated' OR 'calibration'/exp OR 'calibration' OR 'calibrated' OR 'calibrates' OR 'calibrate' OR 'discrimination'/exp OR 'discrimination'
OR 'discriminates' OR 'discriminate' OR 'discriminated' OR 'concordance'/exp OR 'concordance' OR 'c-statistic' OR 'c-index' OR 'roc curve'/exp OR 'roc curve' OR 'roc' OR 'area under curve'/exp OR 'area under curve' OR 'auc'/exp OR 'auc' OR 'development'/exp OR 'development')

NOT
('preterm' OR 'neonatal' OR 'infant, premature'/exp OR 'infant, premature' OR 'premature'/exp OR 'premature' OR 'infant'/exp OR 'infant' OR 'infants'/exp OR 'infants' OR 'infant, newborn'/exp OR 'infant, newborn' OR 'newborn'/exp OR 'newborn' OR 'newborns' OR 'neonate'/exp OR 'neonate' OR 'neonates')

S3 Table Part A. Complete standardized form. Population characteristics of the 28 articles about development models

Publication	Name of the tool	Study design	Location of inclusion	Date of inclusion	Type of injuries	GCS	Population	Inclusion criteria	Non-inclusion criteria	Management of missing data	Main outcome	Secondary outcomes	Sample size	Number of events
Chang et al. 2017 Am J Hypertens [1]		Prospective US monocentric cohort	Extracted by ICD code, managed in ICU	2011/01 - 2015/12	ICH	NA	- Age mean 61.6 SD 14.0 - Female NA	- Spontaneous ICH - Time of adm. NA	- underlying vascular lesions - coagulation disorder	NA	Mortality at discharge		672	Mortality: 162 (24%)
Chuang et al. 2009 Int J Qual Health Care [2]	Simplified ICH score	Taiwan monocentric Registry	Registry of the NICU	2006/01 - 2007/12	ICH	32% GCS 3-8 23% GCS 9-13	- Age mean 60.6 SD 16.7 - Female 30%	- Adm. <24h of the onset - spontaneous ICH	- hemato. (leukemia) or coag. disorder	NA	Mortality at 1 month		NA: 217? 293?	Overall 40/293 (14%)
Di Napoli et al. 2011 Stroke [3]		Prospective Argentine multicenter (2) cohort	Admission in ICU	2005/11 - 2009/12	ICH	median 13 IQR 10-15	- Age mean 67.3 SD 11.5 - Female 42%	- Adm. <24h of the onset - spontaneous ICH	- history of infection, comorbidities, acquired in-hosp infection	- predictors: exclusion (n=19) - outcome: none	Overall mortality at 1 month		210	Mortality 63 (30%)
Edwards et al. 1999 Neurology [4]		Prospective US monocentric cohort	Admission in NICU	1996/12 - 1997/08	ICH	mean 9.8 SD 3.9	- Age mean 62.4 SD 16.1 - Female 42%	- supra-tentorial ICH - Time of adm. NA	- SAH - CT >24h after onset	NA	Mortality at discharge		81	Mortality: 21 (26%)
Fallenius et al. 2019 Stroke [5]		Retrospective analysis of prospective multicenter (4) cohort	Admission in ICU	2003 - 2013	ICH	Median 8 IQR 4-13	- Age median 61 IQR 52-69 - Female NA	Spontaneous ICH	- Isolated IVH	- Predictors: Complete Case only (Exclusion n=53) - Outcome: none	Mortality at 12 months		972	Mortality 421 (43%)
Godoy et al. 2006 Stroke [6]	Modified ICH-score (mICH-A, mICH-B)	Prospective Argentine bicentric cohort	Admission in ICU	2003/01 - 2004/07	ICH	median 11 IQR 7-14	- Age mean 66 SD 12 - Female 37%	- Adm. within <24h after onset - Spontaneous ICH	- brain tumours - haemorrhagic transformation of cerebral infarct - aneurysmal or vascular malformation rupture	- Predictors: exclusion - outcome: NA	Mortality at 1 month	Unfav. outcome (GOS 1-3) at 6 months	153	- Mortality at 1 month: 53 (35%) - Unfav. outcome at 6 months: 94 (62%)
Ho et al. 2016 SpringerPlus [7]		Taiwan prospective monocentric registry	Admission to NICU from ED	2009/01 - 2011/12	ICH	NA	- Age mean 62 SD 15 - Female 38%	- Adm. within <24h after onset - Spontaneous ICH		NA	Mortality at discharge		805	Mortality 164 (20.4%)

Publication	Name of the tool	Study design	Location of inclusion	Date of inclusion	Type of injuries	GCS	Population	Inclusion criteria	Non-inclusion criteria	Management of missing data	Main outcome	Secondary outcomes	Sample size	Number of events
Jeng et al. 2008 J Neurol Sci [8]		Taiwan monocentric Registry	Admission in stroke ICU	2002/11 - 2006/12	ICH	mean 10.6 SD 4.2	- Age mean 61.6 SD 13.5 - Female 39%	- Adm <12h of the onset - (Ischemic stroke or) non-trauma ICH	- Rapid improvement - Transient ischemic attack - Anoxic-ischemic brain injury - SAH	NA	Mortality at 3 months	- Unfav. outcome (mRS>2 or Bartel index <80) at discharge - Mortality at 3 months	342	- Mortality: 62 (18%)
Lukic et al. 2012 Acta Neurol Belg [9]		Prospective data Origin NA	Admission in NICU	2005 - 2009	ICH	GCS mean 9 SD 4	- Age mean 67 SD 11 - Female 52%	Spontaneous ICH, adm <6h of the onset, medically treated	Oral anticoagulant therapy, severe concomitant disease or disability	NA	Mortality at discharge		411	Mortality 256 (62%)
Maas et al. 2017 Cerebrovasc Dis [10]		US monocentric prospective cohort	Admission in neuro-spine ICU	2010/01-2016/03	ICH	GCS Median 14 IQR 8-15	- Age Mean 64.3 SD 13.6 - Female 50%		- Death from withdrawal of care - secondary ICH	Complete case analysis (exclusion n=135)	Unfav. Outcome (mRS 4-6) at 3 months		254	Fav outcome: 122 (48%)
Sembill et al. 2017 Neurology [11]	Max ICH-score	Retrospective analysis on prospective German monocentric registry	Admission in ICU	2007/01 - 2011/12	ICH	Median 13 IQR 10-15	- Age mean 70 SD 12 - Female 45.4%	Maximally treated spontaneous ICH	Early care limitations (<24 hours) (n=71)	- Predictors: exclusion (n=NA) - Outcome: exclusion (n=18)	Mortality at 12 months	- Mortality at 3 months - Unfav outcome (mRS 4-6) at 3 and 12 months	471	- 12-m mortality 142 (30.1%) - 12-m unfav outcome 214 (45.4%)
Ziai et al. 2015 Neurocrit Care [12]		Retrospective US bicentric cohort	Extracted by ICD code + managed in NICU	2003 - 2010	ICH	median 7 IQR 9	- Age mean 61.8 SEM 1.2 - Female 48%	- Spont. IVH - Adm within 24h of the onset	- Aneurysmal SAH - ICH w/ underlying lesions (tumor, AVM, aneurysm)	- Predictors: exclusion (13) - Outcome: NA	Mortality at discharge	Unfav outcome (mRS 4-6) at discharge	170	- Mortality: 87 (51%) - Unfav. outcome: 144 (85%)
Celi et al. 2012 J Pers Med [13]		Retrospective analysis on prospective US monocentric cohort	Admission in ICU	1995/01 - 2006/02	SAH	NA	NA			NA	Mortality at discharge		MIMIC database 150	Mortality: 57 (25.6%)
Claassen et al. 2004 Crit Care Med [14]	SAH-PDS	Prospective US monocentric cohort	Admission in NICU	1996/07 - 2002/06	SAH	NA	- Age mean 54 SD 14 - Female 71%	- Adm. < 3d after onset	- AVM	Predictors: NA Outcome: exclusion (n=NA)	Unfav. outcome (mRS 4–6) at 3 months		413	Unfav outcome: 167 (40.4%)
Czorlich et al. 2015 Acta Neurochir [15]	Improved SAPS II	German monocentric registry	All treated in ICU, recruitment location NA	2010/11 - 2014/11	aSAH	14-15 58% 11-13 8% 9-10 4% 6-8 8% 3-5 22%	- Age mean 54.4 SD 13.74 - Female 70%	- aneurysmal SAH	- Angiogram-negative permesencephalic SAH (54) - AVM - prior syndromic disease	Predictors: exclusion (n=21) Outcome: NA	Mortality at discharge		242	Before exclusion Mortality: 49/263 (18.3%)

Publication	Name of the tool	Study design	Location of inclusion	Date of inclusion	Type of injuries	GCS	Population	Inclusion criteria	Non-inclusion criteria	Management of missing data	Main outcome	Secondary outcomes	Sample size	Number of events
Degos et al. 2012 Anesth [16]		Prospective French monocentric cohort	Admission in NICU	2002/01 - 2010/12	aSAH	median 14 IQR 12-15	- Age <60y: 708 60-70y: 138 ≥70y: 87 - Female 62%	- aneurysmal SAH angiographically confirmed - Treated = coiling or clipping	- No aneurysm procedure (n=67)	Predictors: exclusion (n=21) Outcome: none	Unfav. outcome (mRS 4-6) at 12 months (follow up visits or phone)		933 (526 from Degos et al, 2012 Stroke)	Unfav outcome: 180 (19.3%)
Degos et al. 2012 Stroke [17]	ABC score	Prospective French monocentric cohort	Admission in NICU	2003/01 - 2009/12	aSAH	median 14 IQR 11-15	- Age mean 50 SD 13 - Female 64%	- aneurysm SAH angiographically confirmed - coiled with or without stents	- invasive treatment (n=48) - open surgical clipping (n=168)	Predictors: exclusion (n=10) Outcome: NA	Mortality at 12 months	- independt function (mRS 0-3) at 12 months - full recovery (mRS 0-1) at 12 months	368	Mortality: 64 (17.4%) mRS 0-3: NA mRS 0-1: 257 (69.8%)
Kissoon et al. 2015 J Stroke Cerebrovasc Dis [18]		US monocentric registry	Admission in NICU	2001/10 - 2011/06	aSAH	WFNS: mean 2.3 SD 1.5	- Age mean 55.7 SD 13.5 - Female 66%	Aneurysmal SAH		Predictors: NA Outcome: exclusion (n=19)	Unfav. outcome (mRS 3-6) "during follow up" (mean 8 ± 8 months)		288	Unfav. outcome: 98 (34%)
Konczalla et al. 2016 World Neurosurg [19]		German monocentric registry	Surgical database, all admitted in NICU	2003 - 2012	SAH	WFNS 4-5 57%	- Age mean 53 SD 12 - Female 71%	- long lasting (>14d) cerebral vasospasm - severe cerebral vasospasm or neuro deterioration + moderate-to-severe vasospasm		NA	Fav outcome (mRS 0-2) at 6 months		106	Fav outcome: 64 (60%)
Schuiling et al. 2005 J Neurol Neurosurg Psychiatry [20]		Prospective Dutch monocentric cohort	Admission in ICU	2002/06 - 2004/02	SAH	WFNS 4-5 47%	- Age NA - Female 64%	<24 h after the onset		NA	Unfav outcome (mRS 4-6) at 3 months (follow up visit)		68	Unfav outcome: 40 (59%)
Schuiling et al. 2005 Neurosurgery [21]		Retrospective Dutch monocentric cohort	Admission in ICU	2000/01 - 2002/06	SAH	WFNS 4-5 35%	- Age mean 55 range 17-93 - Female 73%	Adm < 4 d after onset	- non-aneurysmal perimesencephalic hemorrhage - moribond on adm	- Predictors: exclusion (n=2) - Outcome: NA	Unfav outcome (mRS 4-6) at 3 months		136	Unfav outcome: 65 (48%)

Publication	Name of the tool	Study design	Location of inclusion	Date of inclusion	Type of injuries	GCS	Population	Inclusion criteria	Non-inclusion criteria	Management of missing data	Main outcome	Secondary outcomes	Sample size	Number of events
Szklener et al. 2015 BMJ open [22]		Prospective Polish monocentric cohort	Poor grade SAH (WFNS IV-V) disqualified from surgery, admitted in NICU	2001/01 - 2010/12	SAH	WFNS 4 27% WFNS 5 73%	- Age mean 57 range 21-87 - Female 43%	- Non-operated (disqualified) SAH - Adm <24h from the onset	- peri mesencephalic patterns of haemorrhage on CT - intoxication - Prior serious medical conditions	NA	Unfav outcome (mRS 5-6) at 1 month		101	Unfav outcome: 80 (79%)
Weiss et al. 2006 Anesthesiology [23]		France monocentric Cohort	Admission in NICU after surgery	2003/12 - 2004/10	SAH	WFNS 4-5 33%	- Age mean 48 SD 11 - Female 57%	- <2d after onset - evidence of bleeding on CT - aneurysm at angiography	- No surgical or endovascular treatment (n=4) - surgery / coiling > 48h after adm (n=7)	- Predictors: simple imputation by last value (S100B) - Outcome: NA	Unfav outcome (GOS 1-3) at 6 months		74	Unfav outcome: 24 (32%)
Witsch et al. 2016 Ann Neurol [24]		Prospective US monocentric cohort (SHOP)	Admission in ICU	1996/07 - 2014/03	SAH	28% 3-8 10% 9-12 62% 13-15	- Age mean 55.3 SD 14.5 - Female 68%	- Adm < 14d from the onset	- AVM	Predictors: exclusion (n=93) Outcome: multiple Imputation for mRS (n=351) by MCMCM (Little's MCAR test not significant)	Unfav outcome (mRS 4-6) at 12 months (by phone)	- TICS (cognitive status) at 12 months - SIP (QOL - physical) at 12 months	mRS 1526 cog 699 QOL 401	Unfav outcome: 1200 (79%)
Zafar et al. 2017 Neurocrit Care [25]		Retrospective US monocentric cohort	Hosp. database with high grade SAH (≥HH3F3), recruitment location NA	2011/09 - 2016/02	SAH	mean 10.4 SD 4.7	- Age mean 58.3 SD 14.2 - Female 69%	- aneurysmal SAH - high grade H&H≥3		- Predictors: exclusion of variables with >10% missing date (n=22), "imputing for the rest" - Outcome: NA	Mortality at discharge	- Unfav outcome (GOS 1-3) at discharge - functionnal outcome (GOS 1-2, 3, 4-5) at discharge	153	Mortality: 28 (18%)
Zhao et al. 2017 J Neurosurg [26]		Prospective Chinese multicentric (11) cohort	Poor grade SAH (WFNS IV-V) recruited at the ED	2010/10 - 2012/03	SAH	mean 7.5 SD 2.6	- Age mean 54.6 SD 11.8 - Female 47%	- poor grade aSAH WFNS 4-5 - endovascular treatment	- neurological improvement after resuscitation	NA	Unfav outcome (mRS 4-6) at 12 months		136	Unfav outcome: 64 (47%)
Zheng et al. 2019 Front Neurol [27]		Prospective Chinese Multicentric study	Poor grade aSAH (WFNS IV-V)	2010/10 - 2012/03	aSAH	WFNS V 53.6%	- Age Mean 55 SD 11.6 - Female 50.9%	- aneurysm at angiography / MRI		- Predictors and outcome: exclusion	Unfav outcome (mRS 4-6) at 12 months		324	Unfav outcome: 190 (58.6%)

Publication	Name of the tool	Study design	Location of inclusion	Date of inclusion	Type of injuries	GCS	Population	Inclusion criteria	Non-inclusion criteria	Management of missing data	Main outcome	Secondary outcomes	Sample size	Number of events
Weimer et al. 2016 Crit Care Med [28]		Retrospective analysis on prospective US monocentric cohort	Admission in NICU	2008/08 - 2011/10	SAH 35% SDH 35% IPH 30%	- Died: median 7 (IQR 4-10) discharged: 15 (10-15) - mRS 0-3: 15 (12-15) mRS 4-6: 10 (7-14)	- Age: *Discharged median 69 (IQR 53–76) *Died: 62 (51–76) - Female 53%	- Aneurysmal or CT-neg SAH, subdural hematoma or Intra-parenchymal hemorrhage - Time of adm NA	- SAH secondary to vascular dissection - vasculopathy - AVM - Other aneurysmal causes	- Predictors: NA - Outcome at 12m: simple imputation with the outcome at 3m (n=53) - Outcome at 3m: exclusion (n=29)	Mortality at discharge	Unfav outcome (mRS 4-6) at 12 months (phone interview by NA)	Mortality: 357 Unfav outcome: 328	- Mortality: 41 (11%) - Unfav outcome: 156 (48%)

S3 Table - Part B. Complete standardized form. Prognostic tools details of the 28 articles about development models.

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Chang et al. 2017 Am J Hypertens [1]	9 tools Logistic regression	Yes (n=11)	Univariate (p<0.001) then multivariate (NA)	(3 or 4): Hematoma volume, NIHSS + depending on the models: mean PP (dich), mean BP (dich), creatinine, IVH	12 hours after admission	No (OR+CI w/o intercept)	No	AUC ROC: cf table 4	No	No	No	No	No
Chuang et al. 2009 Int J Qual Health Care [2] Simplified ICH score	2 tools Logistic regression then points assigned on the strength of association w/ outcome	Yes (n=8)	Univariate <0.1 then multivariate (forward stepwise p<0.05)	Both (5): age, GCS, ATCD HTA, glc, dialysis dependency	First evaluation	- LR: No (OR + CI w/o intercept) - sICH: yes	No	AUC ROC: *LR: 0.91 *sICH: 0.89 (0.84-0.94)	Accuracy. Se. Sp. PPV. NPV. LR+. LR-: cf table 4 (Cut-off: Youden)	HL GOF test: *LR: p=0.55 *sICH: p=0.34 & histogram obs / pred	ICH score ICH-GS - pairwise comparison of ROC curves - McNemar test to compare Se & Sp	10-fold Cross-Validation	No
Di Napoli et al. 2011 Stroke [3]	3 tools Logistic regression	Yes	None : adding biomarkers to the ICH score (Hemphill et al.)	(6) ICH score + glucose or WBC or CRP	Admission	No (OR+CI w/o intercept)	Nagelkerke R^2 Glc 68.8 WBC 70.7 CRP 71.8 LR chi^2 Glc 174.7 WBC 179.6 CRP 182.2	AUC ROC Glc 0.973 WBC 0.976 CRP 0.978	No	HL GOF test p>0.2	Yes, ICH score - Net benefit decision curve - NRI Glc 3.3% p=0.57 WBC 2.19% p=0.56 CRP 8.14% p=0.6	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Edwards et al. 1999 Neurology [4]	2 tools - Logistic regression - Artificial neural network	Yes - LR: n=8 - ANN: n=14 Interactions	- LR: univ ($p \leq 0.25$) then multivariate $p<0.1$ backward (clinical predictors) then forward (CT predictors) - ANN: NA	*LR (4): Gender, GCS score <=8, CT pineal shift, CT hydrocephalus *ANN (14): age, gender, race, MAP, PP, GCS, history of hypertension, history of diabetes, CT hydrocephalus, CT IVH, CT hematoma size, CT hematoma location, CT cisternal effacement, CT pineal shift	Admission	- LR: No (coef+SE w/o intercept) - ANN: No	No	AUC ROC: *LR: 0.919 *ANN: 0.984	Correct classification rate: *LR: 90% survivors. 79% dead *ANN: 100% both (Cut off: arbitrary probability of 0.5)	HL GOF test: *LR: p: 0.439 *ANN: p: 0.995	Tuhrim equation	No	No
Fallenius et al. 2019 Stroke [5]	3 tools – Logistic regression	No	known prognostic factors from the literature and significant variables from univariate analyses	*Clinical (4) age, GCS, severe chronic comorbidity, modified SAPS II *CT (4) brain stem ICH, hematoma volume, midline shift, IVH *Clinical + CT (7) age, GCS, severe chronic comorbidity, modified SAPS II, brain stem ICH, hematoma volume, IVH	24 hours after admission	No (OR+CI w/o intercept)	Nagelkerke R^2 *Clinical 0.42 *CT 0.22 *Clinical+CT 0.47	AUC ROC *Clinical 0.83 (0.81-0.86) *CT 0.73 (0.70-0.77) *Clinical + CT 0.85 (0.83-0.88)	No	HL GOF test P>0.05	No	No	No
Godoy et al. 2006 Stroke [6] Modified ICH-score	4 tools - analysis NA	Yes (n=5)	None (change of cut offs and one variable removed from ICH score)	*Model A (5): GCS, ICH volume, presence of IVH (depending on Graeb's score), age, comorbidities *Model B (5): same [diff btw the 2 scores = cut offs of GCS, Graeb's score and age]	72 hours after admission	No	No	Non-param. AUC ROC: *30-Day mortality -A: 0.878 (0.824-0.9931) -B: 0.869 (0.811-0.928) *6-month GOS -A: 0.893 (0.844-0.941) -B: 0.895 (0.847-0.943)	Se. Sp. PPV. NPV: cf table 3 (Cut-off: Youden)	No	ICH-score Comparing AUC	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Ho et al. 2016 SpringerPlus [7]	1 score Logistic regression → Nomogram	Yes (n=9)	Univariate (p<0.05) then forward selection	(6) Age, gender, adm NIHSS, systolic BP, Heart disease history, Creatinine	Admission	No (OR+CI w/o intercept)	No	AUC ROC 0.87	No	- Calibration curve - le Cessie and Houwelingen GOF test (p=0.36)	No	No	No
Jeng et al. 2008 J Neurol Sci [8]	2 tools - Cox regression (mortality) - Logistic regression (func outcome)	Yes (n=22) Interactions tested	Univariate (p<0.1) then multivariate (NA)	(6): age, BMI, NIHS, requiring ventilator aid, ICH volume >=30, ventricular extension	Admission	No (HR/OR +CI w/o intercept)	R^2: *Cox (mortality): 68.2% *LR (func outcome): 64.1%	AUC ROC: *Cox (mortality): 0.961 (0.936-0.985) *LR (poor outcome): 0.903 (0.866-0.940)	No	No	No	No	No
Lukic et al. 2012 Acta Neurol Belg [9]	2 tools - Logistic regression - Artificial neural network	Yes (n=8)	- LR: univ (p≤0.20) then multivariate backward (selection NA) - ANN: trial-and-error process	- LR (5): level of consciousness (4 cat), gender, age, pulse BP, verbal GCS - ANN (8): age, gender, pulse BP, mean BP, eye GCS, motor GCS, verbal GCS, level of counsciousness	Admission	- LR: No (coef + SE w/o intercept) - ANN: No	No	AUC ROC LR: 0.86 (0.82-0.89) ANN: 0.94 (0.85-0.99)	ANN on internal validation: True - 90.5% True + 95.1%	HL GOF test LR: p=0.2 ANN: p=0.6	No	Yes for ANN 62 patients (for classification only)	No
Maas et al. 2017 Cerebrovasc Dis [10]	1 tools logistic regression	No Interactions tested	backward conditional selection (elimination based on change in the likelihood ratio)	5 Age, premorbid mRS, IVH by day 5, hispanic ethnicity, GCS by day 5	Day 5	No (OR+CI w/o intercept)	Nagelkerke R^2 0.46 -2 log likelihood 148.1	No	PPV 79.1% NPV 87.1% Diagnostic effectiveness 83%	No	Yes (ICH score)	No	No
Sembill et al. 2017 Neurology [11] Max ICH-score	2 tools (logistic regression) → score	No	NA	(6) lobar ICH vol, non-lobar ICH vol, age, NIHSS, IVH, oral anticoag	24 hours after admission	Yes	No	AUC ROC - mRS 12m: 0.81 (0.77-0.85) - mortality 12m: 0.77 (0.72-0.81)	No	Histogram mRS vs max ICH score	Yes, ICH and MICH score, method by Hanley and McNeil	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Ziai et al. 2015 Neuro Crit Care [12]	4 tools Logistic regression	Yes (n=6)	univariate ($p \leq 0.1$) then multivariate ($p < 0.2$) backwards stepwise then AIC	*mortality, full (4): 4-pt TIL score, ICH vol, IVH vol, DNR at 24h *mortality, ICH score (3): 3-pt TIL score, ICH score, DNR at 24h *mRS, full (4): 4-pt TIL score, GCS, age, ICH vol *mRS, ICH score (2): 4-pt TIL, ICH score	First 72 hours after admission	No (OR+CI w/o intercept)	No	AUC ROC (on internal validation) *mortality, full: 0.94 *mortality, ICH score: 0.96 *mRS, full: 0.94 *mRS, ICH score: 0.92	No	No	No	3-fold Cross-Validation	No
Celi et al. 2012 J Pers Med [13]	3 tools LR; BN; ANN	Yes (n=13)	correlation based feature subset algorithm	(12): Age, Gly, gly SD, max WBC, INR, min GCS, max GCS, mean GCS, min sBP, min NA, mean Na, SD Na	24 hours after admission	No (estimate coef and SE w/o intercept)	No	*AUC ROC LR 0.945; BN 0.958; ANN 0.868 *Mean absolute error LR 0.158; BN 0.127; ANN 0.168	Accuracy LR 89%; BN 87.7%; ANN 83.6%	HL GOF test (LR only): p=0.516	SAPS (factual)	Random Split N=73 No perf reported	No
Claassen et al. 2004 Crit Care Med [14] SAH PDS	1 tool Logistic regression then score based on the weight of each coeff of the LR	Yes (interaction) (n=NA)	univariate then multivariate forward stepwise	(4): arterio-alveolar gradient of >125 mm Hg, HCO3 of <20 mmol/L, Glucose of >180 mg/dL, mean arterial pressure of <70 or >130 mm Hg	24 hours after admission	LR: No (OR+CI w/o intercept) Score: yes	No	AUC ROC 0.79 (0.74 – 0.85)	No	Plot	APACHE-II SIRS summary score SAH sum score (comparing AUC)	No	No
Czorlich et al. 2015 Acta Neurochir [15]	1 tool Logistic regression	Yes (n=NA)	univariate ($p<0.1$) then multivariate forward	(3): SAPS-II, anticoag drugs, headache	24 hours after admission	No (OR+CI w/o intercept)	No	AUC ROC 0.860 (0.786-0.934)	No	No	SAPS-II (comparing AUC)	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Degos et al. 2012 Anesth [16]	2 tools Logistic regression	Yes (interaction) n=14	univ p<0.2 then multivariate (backward and forward) + interaction	*w/o interaction (9): Intracranial hypertension (IH) on adm, severe IH, isch vasospasm, rebleeding, endovascular complication, surgery complication, Fisher score III-V, admision hydrocephalus, >60y *interaction: same + hydrocephalus*age > 60y	NA (neurologic events recorded during the ICU stay)	No (OR+CI w/o intercept)	No	AUC ROC (dev? IV?) *interaction: 0.85 (0.82-0.88) *w/o interaction: 0.84 (0.82-0.88)	No	HL GOF test (dev? IV?) *interaction : p=0.22 *w/o interaction: p=0.18	No	jacknife bootstrap 100 iterations	No
Degos et al. 2012 Stroke [17] ABC score	3 tools Logistic regression then score based on the weight of each ORs of the LR	Yes (n=9)	uni puis multi stepwise - p - most parsimonious model	(3): troponin I, S100B, GCS	Admission	LR: No (OR+CI w/o intercept) Score: yes	No	AUC ROC *mortality : 0.828 (0.772-0.885) *full recovery 0.83 (0.79-0.88) *independant : 0.82 (0.77-0.88)	No	HL GOF test NA	WFNS score Fisher score - IDI - NRI -risk stratification capacity (supl met)	Temporal 2008-2009 N= 158 mortality : 0.76 (0.67-0.86) Independent : 0.76 (0.67-0.86)	No
Kissoon et al. 2015 J Stroke Cerebrovasc Dis [18]	2 tools Logistic regression	Yes (n=NA)	univariate	*Model 1 (5): positive fluid balance, WFNS, transfusion, glc, cerebral infarction *Model 2 (5): Model 1 + propensity score	End of NICU stay (fluid balance)	No (OR+CI w/o intercept)	No	AUC ROC *1: 0.91 *2: 0.92	No	No	No	No	No
Konczalla et al. 2016 World Neurosurg [19]	1 tool Logistic regression	Yes (n=3)	univariate then multivariatep<0.05	(3): age <55y, admission WFNS I-III, small ICH	Admission	No (OR+CI w/o intercept)	Nagelkerke R^2 0.267	No	No	No	No	No	No
Schuiling et al. 2005 J Neurol Neurosurg Psychiatry [20]	2 tools Logistic regression	Yes n=5	Univariate p<0.1 then multivariate forward selection	*Model 1 (3): WFNS, age, Hijdra score *Model 2 (4): Model 1 + troponin I	24 hours after admission	No (OR+CI w/o intercept)	No	AUC ROC *w/o troponin 0.86 (0.77 - 0.95) *w/ troponin 0.89 (0.81 - 0.97)	No	No	No	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Schuiling et al. 2005 Neurosurgery [21]	1 tool Logistic regression	Yes (n= 4)	Univariate then multivariate forward selection p<0.1	(2) WFNS, Hijdra score	24 hours after admission	No (OR+CI w/o intercept)	No	AUC ROC 0.81 (0.73 - 0.88)	No	No	SAPS II (comparing AUC)	No	No
Szklener et al. 2015 BMJ open [22]	1 tool Logistic regression then grading system	Yes (n=5)	univariate then multivariate backward p<0.05	(4): WFNS, age, Fisher scale, leucocytosis	First hours after hospital admission	No (OR+CI w/o intercept)	No	AUC ROC (grading scale only): 0.91	No	HL GOF test (LR only) p=0.9322	No	No	No
Weiss et al. 2006 Anesthesiology [23]	1 tool Logistic regression	Yes (n=6)	Univariate (p<0.2) then multivariate	(3): age, WFNS score, Mean daily S100B>0,4 g/l	8 days after admission	No (OR+CI w/o intercept)	No	AUC ROC 0.88 (0.8-0.96)	No	HL GOF test 0.84	No	No	No
Witsch et al. 2016 Ann Neurol [24] FRESH score	3 tools Linear regression	Yes n=35	mix of knowledge-based and data-driven approaches (BIC k-means)	*FRESH (4): Hunt&Hess, APACHE w/o GCS, age, aneurysmal rebleed *FRESH-cog (5): FRESH + education *FRESH-quol (6): FRESH + education + premorbid disabilities	48 hours after admission	Yes	Nagelkerke R^2 and Cox/Snell R^2 (dev? IV?) *FRESH: Nagelkerke R^2 0.50 Cox/Snell R^2 0.35 *FRESH -cog & -quol: NA	AUC ROC (dev? IV?) *FRESH: 89.8% (88.1-91.6) *FRESH-cog: 79.7 (75.2-84.2) *FRESH quol: 78.2 (71.3-85.2)	No	No	HAIR Delong et al method (AUCs)	nonparametric bootstrap using 500 repetitions	CONSCIOUS-1 (52 centres) N= 413 N-R^2 0.2 ; C/S R^2 0.13 AUC ROC: 73.2 (67.3–79.1)
Zafar et al. 2017 Neurocrit Care [25]	3 tools Logistic regression	Yes (n=451)	Multivariate (Lasso penalty and bootstrapping)	*Mortality (3): APACHE II, glucose, ICP *GOS 1-3 (2): Leveciteram - MV *multilevel (NA): max GCS day 1, min GCS day 2-3, APACHE II	72 hours after admission	No	No	AUC ROC for binary models *mortality: 0.9198 *GOS 1-3: 0.9456	No	Yes (multilevel model only): Bar plot	No	Cross validation	No
Zhao et al. 2017 J Neurosurg [26] AMPAS	2 tools Logistic regression	Yes (n=10)	Univariate (p<0.05). backward multivariate selection	*Pre op (4): age, WFNS, Fisher, wider neck aneurysm *Post op (5): pre op + pneumonia	NA *pre-op: median 24h range 0-35 days *post op: "during the ICU stay"	No (OR+CI w/o intercept)	No	AUC ROC *pre op: 0.86 (0.80-0.92) *post op: 0.87 (0.81-0.93)	No	HL GOF test *pre op : p=0.941 *post op : p=0.653	No	No	No

Publication	Main statistical analysis	Specified candidate predictors	Predictor's selection strategy	Predictors (number)	Landmark time	Public equation?	Global performances	Reporting of discrimination	Reporting of Classification	Reporting of calibration	Comparing with existing score	Internal validation	External validation
Zheng et al. 2019 Front Neurol [27]	5 tools: 4 Logistic regressions 1 score (WAP)	Yes (n=22)	backward multivariate selection	- Model 1 (3): age, ventilated y/n, pupil react - Model 2 (3): age, pupil react, GCS - Model 3 (4): age, pupil react, GCS, mFisher - Model 4 (5): age, pupil react, GCS, mFisher, ttt modality - WAP score (3): WFNS, age, pupillary reactivity	3 days	No for the models (OR+CI w/o intercept)	No	AUC ROC - M1: 0.74 (0.69-0.79) - M2: 0.81 (0.76-0.86) - M3: 0.85 (0.81-0.89) - M4: 0.86 (0.82-0.90) - WAP score: 0.77 (0.72-0.82)	No	WAP score only: - HL GOF test p=1.00 - Table obs vs pred	No	No	No
Weimer et al. 2016 Crit Care Med [28]	2 tools Logistic regression	Yes (n=NA)	Multivariate (backward selection on p>0.05)	*Mortality (6): GCS, no surg intervention, vasopressor use, renal failure, hist of CV disease, history of BPCO *mRS (9): age, NIHSS, brainstem herniation, type of bleed, arrhythmia, premorbid mRS, hist of diabetes, hist of cancer, hist of BPCO	NA (some variables recorded during the ICU stay)	Yes	No	C-stat: *Mortality: 0.96 *mRS: 0.92	No	HL GOF test: *Mortality: p=0.98 *mRS: p=0.95	No	No	No

References of the S3 Table

1 Chang JJ, Khorchid Y, Dillard K, *et al.* Elevated Pulse Pressure Levels Are Associated With Increased In-Hospital Mortality in Acute Spontaneous Intracerebral Hemorrhage. *Am J Hypertens* 2017;**30**:719–27. doi:10.1093/ajh/hpx025

2 Chuang YC, Chen YM, Peng SK, *et al.* Risk stratification for predicting 30-day mortality of intracerebral hemorrhage. *Int J Qual Health Care J Int Soc Qual Health Care* 2009;**21**:441–7. doi:10.1093/intqhc/mzp041

3 Di Napoli M, Godoy DA, Campi V, *et al.* C-reactive protein level measurement improves mortality prediction when added to the spontaneous intracerebral hemorrhage score. *Stroke* 2011;**42**:1230–6. doi:10.1161/STROKEAHA.110.604983

4 Edwards DF, Hollingsworth H, Zazulia AR, *et al.* Artificial neural networks improve the prediction of mortality in intracerebral hemorrhage. *Neurology* 1999;**53**:351–7.

5 Fallenius M, Skrifvars MB, Reinikainen M, *et al.* Spontaneous Intracerebral Hemorrhage. *Stroke* 2019;**50**:2336–43. doi:10.1161/STROKEAHA.118.024560

6 Godoy DA, Pinero G, Di Napoli M. Predicting mortality in spontaneous intracerebral hemorrhage: can modification to original score improve the prediction? *Stroke* 2006;**37**:1038–44. doi:10.1161/01.str.0000206441.79646.49

7 Ho W-M, Lin J-R, Wang H-H, *et al.* Prediction of in-hospital stroke mortality in critical care unit. *SpringerPlus* 2016;**5**:1051. doi:10.1186/s40064-016-2687-2

8 Jeng JS, Huang SJ, Tang SC, *et al.* Predictors of survival and functional outcome in acute stroke patients admitted to the stroke intensive care unit. *J Neurol Sci* 2008;**270**:60–6. doi:10.1016/j.jns.2008.01.015

9 Lukić S, Ćojbasić Ž, Perić Z, *et al.* Artificial neural networks based early clinical prediction of mortality after spontaneous intracerebral hemorrhage. *Acta Neurol Belg* 2012;**112**:375–82. doi:10.1007/s13760-012-0093-2

10 Maas MB, Francis BA, Sangha RS, *et al.* Refining Prognosis for Intracerebral Hemorrhage by Early Reassessment. *Cerebrovasc Dis* 2017;**43**:110–6. doi:10.1159/000452679

11 Sembill JA, Gerner ST, Volbers B, *et al.* Severity assessment in maximally treated ICH patients: The max-ICH score. *Neurology* 2017;**89**:423–31. doi:10.1212/WNL.0000000000004174

12 Ziai WC, Siddiqui AA, Ullman N, *et al.* Early Therapy Intensity Level (TIL) Predicts Mortality in Spontaneous Intracerebral Hemorrhage. *Neurocrit Care* 2015;**23**:188–97. doi:10.1007/s12028-015-0150-9

13 Celi LA, Galvin S, Davidzon G, *et al.* A Database-driven Decision Support System: Customized Mortality Prediction. *J Pers Med* 2012;**2**:138–48. doi:10.3390/jpm2040138

14 Claassen J, Vu A, Kreiter KT, *et al.* Effect of acute physiologic derangements on outcome after subarachnoid hemorrhage. *Crit Care Med* 2004;**32**:832–8.

15 Czorlich P, Sauvigny T, Ricklefs F, *et al.* The simplified acute physiology score II to predict hospital mortality in aneurysmal subarachnoid hemorrhage. *Acta Neurochir (Wien)* 2015;**157**:2051–9. doi:10.1007/s00701-015-2605-3

16 Degos V, Gourraud PA, Tursis VT, *et al.* Elderly age as a prognostic marker of 1-year poor outcome for subarachnoid hemorrhage patients through its interaction with admission hydrocephalus. *Anesthesiology* 2012;**117**:1289–99. doi:10.1097/ALN.0b013e318267395b

17 Degos V, Apfel CC, Sanchez P, *et al.* An admission bioclinical score to predict 1-year outcomes in patients undergoing aneurysm coiling. *Stroke* 2012;**43**:1253–9. doi:10.1161/strokeaha.111.638197

18 Kissoon NR, Mandrekar JN, Fugate JE, *et al.* Positive Fluid Balance Is Associated With Poor Outcomes in Subarachnoid Hemorrhage. *J Stroke Cerebrovasc Dis Off J Natl Stroke Assoc* 2015;**24**:2245–51. doi:10.1016/j.jstrokecerebrovasdis.2015.05.027

19 Konczalla J, Brawanski N, Bruder M, *et al.* Outcome of Patients with Long-Lasting Cerebral Vasospasm After Subarachnoid Hemorrhage: Is Prolonged Treatment for Cerebral Vasospasm Worthwhile? A Matched-Pair Analysis. *World Neurosurg* 2016;**88**:488–96. doi:10.1016/j.wneu.2015.10.033

20 Schuiling WJ, Dennesen PJ, Tans JT, *et al.* Troponin I in predicting cardiac or pulmonary complications and outcome in subarachnoid haemorrhage. *J Neurol Neurosurg Psychiatry* 2005;**76**:1565–9. doi:10.1136/jnnp.2004.060913

21 Schuiling WJ, de Weerd AW, Dennesen PJ, *et al.* The simplified acute physiology score to predict outcome in patients with subarachnoid hemorrhage. *Neurosurgery* 2005;**57**:230–6; discussion 230-6. doi:10.1227/01.NEU.0000166536.42876.9C

22 Szklener S, Melges A, Korchut A, *et al.* Predictive model for patients with poor-grade subarachnoid haemorrhage in 30-day observation: a 9-year cohort study. *BMJ Open* 2015;**5**:e007795. doi:10.1136/bmjopen-2015-007795

23 Weiss N, Sanchez-Pena P, Roche S, *et al.* Prognosis value of plasma S100B protein levels after subarachnoid aneurysmal hemorrhage. *Anesthesiology* 2006;**104**:658–66.

24 Witsch J, Frey H-P, Patel S, *et al.* Prognostication of Long-Term Outcomes after Subarachnoid Hemorrhage: The FRESH Score. *Ann Neurol* 2016;**80**:46–58. doi:10.1002/ana.24675

25 Zafar SF, Postma EN, Biswal S, *et al.* Electronic Health Data Predict Outcomes After Aneurysmal Subarachnoid Hemorrhage. *Neurocrit Care* Published Online First: 5 October 2017. doi:10.1007/s12028-017-0466-8

26 Zhao B, Yang H, Zheng K, *et al.* Preoperative and postoperative predictors of long-term outcome after endovascular treatment of poor-grade aneurysmal subarachnoid hemorrhage. *J Neurosurg* 2017;**126**:1764–71. doi:10.3171/2016.4.jns152587

27 Zheng K, Zhong M, Zhao B, *et al.* Poor-Grade Aneurysmal Subarachnoid Hemorrhage: Risk Factors Affecting Clinical Outcomes in Intracranial Aneurysm Patients in a Multi-Center Study. *Front Neurol* 2019;**10**:123. doi:10.3389/fneur.2019.00123

28 Weimer JM, Nowacki AS, Frontera JA. Withdrawal of Life-Sustaining Therapy in Patients With Intracranial Hemorrhage: Self-Fulfilling Prophecy or Accurate Prediction of Outcome? *Crit Care Med* 2016;**44**:1161–72. doi:10.1097/CCM.0000000000001570

S4 Table - Part A. Complete standardized form. Original publications of the tools externally validated in 25 articles.

Score	First publication	Name of the cohort Country (nb of centres)	Date of inclusion	Initial type of injury	Initial main outcome	Type of tool	Predictors	Public equation/ score?	External validation
Tuhrim equation	Tuhrim et al. 1991 Ann Neurol	Pilot Stroke Data Bank USA (4)	NA	ICH 100%	- Mortality at 1 month - Mortality or Bartel index > 60 at 1 year	Continuous score	4 Pulse pressure, GCS, ICH volume, IVH, IVH*GCS	Yes	Edwards et al. 1999 Neurology (1)
ICH score	Hemphill et al. 2001 Stroke	San Francisco USA (2)	1997 - 1998	ICH 100%	Mortality at 1 month	Risk stratification scale based on the strength of association with outcome from LR model	5 Age (2 cat), GCS (3 cat), ICH volume (2 cat), IVH (y/n), Infra-tentorial origin (y/n)	Yes	Barbieri et al. 2009 J Eval Clin Pract (2) Chuang et al. 2009 Int J Qual Health Care (3) Di Napoli et al. 2011 Stroke (4) Godoy et al. 2006 Stroke (5) Huang et al. 2012 Eur J Neurol (6) Maas et al. 2017 Cerebrovasc Dis (7) Naval et al. 2009 Neurol Res (8) Patriota et al. 2009 Arq Neuropsiquiatr (9) Rodriguez-Fernandez et al 2018 BMJ open (10) Schmidt et al 2018 Neurology (11) Sembill et al 2017 Neurology (12)
ICH-GS	Ruiz-Sandoval et al. 2007 Stroke	Mexico (1)	1999 - 2003	ICH 100%	- Mortality at discharge - Mortality at 1 month	Score (5-13)	5 age (3 cat), GCS (3 cat), ICH location (2 cat), IVH volume (3	Yes	Chuang et al. 2009 Int J Qual Health Care (3)

Score	First publication	Name of the cohort Country (nb of centres)	Date of inclusion	Initial type of injury	Initial main outcome	Type of tool	Predictors	Public equation/ score?	External validation
					- Fav outcome (GOS 4-5) at 1 month		cat), extension into ventricles (y/n)		Naval et al. 2009 Neurol Res (8)
Modified Intracerebral Hemorrhage Score (MICH)	Cho et al. 2008 Crit Care Med	Taiwan (1)	2001 - 2005	ICH 100%	- Mortality at 6 month - Fav. outcome (GOS 4-5) at 12 months - Barthel index (≥ 55) at 12 months	Score (0-5)	3 GCS (3 cat), ICH volume (3 cat), IVH or hydrocephalus (y/n)	Yes	Sembill et al 2017 Neurology (12)
Max ICH score	Sembill et al 2017 Neurology	Germany (monocentric)	2007/01 - 2011/12	Maximally treated ICH 100%	- Mortality at 3 and 12 months - Unfav outcome (mRS 4-6) at 3 and 12 months	Continuous score	6 lobar ICH vol (2 cat), non-lobar ICH vol (2 cat), age (4 cat), NIHSS (4 cat), IVH (y/n), oral anticoag (y/n)	Yes	Schmidt et al. 2018 Neurology (11)
ISAT	Risselada et al. 2010 Eur J Epidemiol	ISAT Europe (RCT - multicentre)	NA	aneurysmal SAH 100%	Mortality at 60 days	Continuous score	4 age (in decades), lumen size (num), Fisher grade (4 cat), and WFNS grade (5 cat + NA)	Yes	Dijkland et al. 2016 Crit Care Med (13)
VASOGRADE	de Oliveira Manoel et al. 2015 Stroke	International 3 SAHIT trials (CONSCIOUS-1, EPO trial, statin trial) + 1 centre (Canada)	NA	SAH 100%	Delayed Cerebral Ischemia	3- cat grading system (green - yellow - red)	2 WFNS (3 cat), modified Fisher Scale (2 cat)	Yes	Dengler et al. 2017 Eur J Neurol (14)
SAH sum score / Hijdra score	Hijdra et al. 1988 Stroke	European (Rotterdam, Amsterdam, Glasgow, London) (Multicenter)	1977-1983	aneurysmal SAH 100%	- Unfav outcome (GOS 1-3) at 4 weeks- DCI - rebleeding	Continuous score	2 SAH volume (num), GCS (num)	Yes	Claassen et al. 2004 Crit Care Med (15)
SAHIT	Jaja et al. 2018 BMJ	SAHIT International (7 RCT = Van den Bergh 2005, IMASH, COUNSCIOUS-1, ISAT, IHAST, MAPS, Etminan 2013 + 2 registries HELBOK 2013, Smith 2005, Reilly 2004 - multicenter)	NA	SAH 100%	- Mortality at 3 months - Unfav outcome (GOS 1-3) at 3 months	Logistic regression	- Core 3: age, premorbid history of hypertension, WFNS on adm - Neuroimaging 6: core + CT vol of SAH, aneurysmal size, aneurysm location - Full 7: neuroimaging + treatment modality	Online calculator	Mascitelli et al 2018 Neurosurgey (16)
HAIR / SAH score	Lee et al. 2014 Neurocrit Care	Chicago USA (1)	2006 - 2011	SAH 100% Exclusion of CT negative	In-hospital mortality	risk stratification scale based on	4 Hunt and Hess grade (3 cat), age (3 cat),	Yes	Witsch et al. 2016 Ann Neurol (17)

Score	First publication	Name of the cohort Country (nb of centres)	Date of inclusion	Initial type of injury	Initial main outcome	Type of tool	Predictors	Public equation/ score?	External validation
						the strength of association of the predictors with the outcome (0-8)	IVH (y/n), re-bleeding within 24 hours (y/n)		Dengler et al. 2017 Eur J Neurol (14) Abulhasan et al. 2017 Neurocrit Care (18)
APACHE II	Knaus et al. 1985 Crit Care Med	USA (13)	1979 - 1982	All admission in ICU	Hospital mortality	Scale (0 - 71)	14 age, t°, mean BP, HR, RR, O2, art pH, Na, K, creat, Haematocrit, WBC, GCS, chronic health point	Yes	Claassen et al. 2004 Crit Care Med (15) Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19) Huang et al. 2012 Eur J Neurol (6) Moon et al. 2015 J Clin Neurosci (20) Rodriguez-Fernandez et al. 2018 BMJ Open (10)
SOFA	Vincent et al. 1996 Intensive Care Med	Expert meeting	NA	Admission in ICU for sepsis	Hospital mortality	Scale (6-24)	6 PaO2/FiO2, platelets, bilirubin, hypotension, GCS, creatinine	Yes	Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19) Basile-Filho et al. 2018 Medicine (21)
SIRS summary score	Bone et al. 1992 Chest	USA (40)	NA	Admission in ICU for sepsis	Hospital mortality	NA	4 t°, HR, RR, WBC count	NA	Claassen et al. 2004 Crit Care Med (15)
SAPS	Le Gall et al 1984 Crit Care Med	France (8)	NA	Admission in ICU	Mortality at ICU discharge	Scale (0-50)	14 Age, HR, sBP, t°, RR or MV, urinary output, urea, haematocrit, WBC count, glc, K, Na, HCO3, GCS	Yes	Handschu et al. 2005 J Neurol (22)
SAPS II	Le Gall et al 1993 JAMA	International (multicentre)	1991 - 1992	Admission in ICU	Mortality at ICU discharge	Scale (0-163)	17 age, HR, sBP, t°, MV, urinary output, urea, WBC count, K, Na, HCO3-, bilirubin, GCS, type of	Yes	Schuiling et al. 2005 Neurosurgery (23) Handshu et al. 2005 J Neurol (22)

Score	First publication	Name of the cohort Country (nb of centres)	Date of inclusion	Initial type of injury	Initial main outcome	Type of tool	Predictors	Public equation/ score?	External validation
							admission, AIDS, hemato malignancy, metastatic cancer		Celi et al. 2012 J Pers Med (24)
									Czorlich et al. 2015 Acta Neurochir (Wien) (25)
									Moon et al. 2015 J Clin Neurosci (20)
									Huang et al. 2012 Eur J Neurol (6)
									Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19)
									Barbieri et al. 2009 J Eval Clin Pract (2)
SAPS III	Moreno et al. 2005 Intensive Care Med	SAPS 3 project European multicentric cohort (303)	2002	All admission in ICU	Mortality at discharge	Score (5-124)	(20) Age, comorbidities, LOS before ICU, location before ICU, use of vasoactive drugs before ICU, planned ICU adm, reasons for ICU adm, surgical status, anatomical site, acute infection at adm, GCS, bilirubine, temperature, creatinine, HR, WBC, pH, platelets, systolic BP, oxygenation	Yes	Basile-Filho et al. 2018 Medicine (21)

S4 Table - Part B. Complete standardized form. Details of the 25 articles about stand-alone external validation studies.

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
Tuhrim equation	Edwards et al. 1999 Neurology (1)	Prospective cohort USA (1) 1996-1997	Supra-tentorial ICH (exclusion SAH)	Mortality at discharge	81 mortality 21 (26%)	- Age Mean 62.4 SD 16.1 - Female 42%	No	No	correct prediction survivor = 88% death 62% (Cut off NA)	No	No
ICH score	Barbieri et al. 2009 J Eval Clin Pract (2)	Registry of ICU Italy (1) period NA	spontaneous intra parenchymal hemorrhage	Mortality at 1 month	81 mortality 49 (60.5%)	- Age Mean 64.6 SD 13.8 - Female: 36%	No	AUC ROC 0.732 (0.617-0.847)	No	No	No
	Chuang et al. 2009 Int J Qual Health Care (3)	Retrospective cohort Taiwan (1) 2006-2007	spontaneous ICH	Mortality at 1 month	293 Mortality 40 (14%)	- Age Mean 60.6 SD 16.7 - Female 30%	accuracy: 74,1%	AUC ROC: 0.74 (0.65-0.83)	Cut off NA (highest Youden) Se, Sp, PPV, NPV, LR+, LR- (table 4)	HL GOF test p>0.05	No
	Di Napoli et al. 2011 Stroke (4)	Prospective Argentine multicenter cohort (2) 2005 – 2009	Spontaneous ICH	Mortality at 1 month	210 Mortality 63 (30%)	- Age mean 67.3 SD 11.5 - Female 42%	Nagelkerke R^2 56.9 LR chi^2 144.7	AUC ROC 0.94	No	HL GOF test P=0.9	No
	Godoy et al. 2006 Stroke (5)	Prospective cohort Argentina (2) 2003-2004	spontaneous ICH	Mortality at 1 month / Unfav outcome (GOS 1-3) at 6 months	153 mortality 53 (35%)	- Age Mean 66 SD 12 - Female 37%	No	AUC ROC 1-m mortality 0.882 (0.830 - 0.934) 6-m GOS 0.844 (0.781 - 0.907)	PPV, NPV ICH scores of 1, 2, 3, and 4 were 2.9%, 30.8%, 61.1%, and 88.2%	No	No
	Huang et al. 2012 Eur J Neurol (6)	Registry of NICU Chine (1) 2000-2011	primary pontine hemorrhage	Mortality at 1 month	75 mortality 31 (41%)	- Age Mean 54.8 SD 12.7 - Female 21%	No	AUC ROC 0.844 (0.757 - 0.931)	Cut off 1.5 (Youden) Se 96.8% Sp 54.5%	HL GOF test p=0.176 Table obs / pred	No
	Maas et al. 2017 Cerebrovasc Dis (7)	Prospective NICU cohort US (1) 2010 - 2016	Spontaneous ICH	Unfav outcome (mRS 4-6) at 3 months	254 Good outcome 122 (48%)	- Age Mean 64.3 SD 13.6 - Female 50%	Nagelkerke R^2 0.36 –2 log likelihood 270.7	No	PPV 66.7% NPV 83.2% Diagnostic effectiveness 73%	No	No
	Naval et al. 2009 Neurol Res (8)	Registry of NICU Baltimore USA (1) 1999-2006	Supra-tentorial, spontaneous ICH Excl of prior mRS 2-5	Mortality at 1 month	125 mortality 29 (23%)	- Age median 63.5 range 34-90 - Female 42%	No	No	Cut off ≥ 3 (Proba 50/50) PPV: 71% NPV 97.7% Se 93.1% Sp 88.5%	chi square test obs/pred p=0,14	No
	Patriota et al. 2009 Arq Neuropsiquiatr (9)	Prospective cohort admission ICU Brazil (1) 2006	spontaneous ICH	Mortality at 1 month / Fav outcome (GOS 4-5) at 12 months	37 1-m mortality 38% 1-y GOS 4-5 : 38%	- Age Mean 67.7 SD 11.2 - Female 51%	No	AUC ROC mortality 0.804 (0.65 - 0.95) GOS 4-5 0.77 (0.60 - 0.89)	mortality cut off: 3 Se 85.7% Sp 65.2% GOS 4-5 cut off ≤2 Se 100% Sp 42%	Histogram obs / pred	No

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
	Rodriguez-Fernandez et al 2018 BMJ open (10)	Prospective multicenter Spanish Cohort admission ICU (3) 2009-2012	Spontaneous ICH	Mortality at 1 month	336 Mortality 176 (52%)	- Age Median 62 IQR 50-70 - Female NA	No	AUC ROC 0.74 (0.69-0.79)	No	- HL GOF test p<0.001 - GiViTI calibration belt p<0.001	No
	Schmidt et al 2018 Neurology (11)	Prospective cohort Chicago US (1) 2010-2017	Spontaneous ICH	Unfav outcome (mRS 4-6) at 3 months	372 Mortality at 3m: 153 (41%) Unfav outcome at 3m: 236 (63%)	- Age Mean 67 SD 14 - Female 51%	Likelihood ratio X^2 p<0.001	AUC ROC 3m mortality 0.83 (0.79-0.88) 3m unfav outcome 0.85 (0.81-0.89)	No	- Histogram score vs observed mortality and poor outcome - Unweighted sum of squared errors test for GOF p>0.3	No
	Sembill et al 2017 Neurology (12)	Prospective cohort Germany (1) 2007-2011	Spontaneous ICH maximally treated	Mortality at 3 and 12 months Unfav outcome (mRS 4-6) at 3 and 12 months	471 12-m mortality 142 (30.1%) 12-m unfav outcome 214 (45.4%)	- Age Mean 70 SD 12 - Female 45.4%	No	AUC ROC 12m mortality 0.69 (0.64-0.74) 12m unfav outcome 0.72 (0.67-0.76)	3m mortality: PPV 37% 12m unfav outcome: PPV 74.7% Cutoff Youden	Histogram ICH score vs observed mortality	No
MICH	Sembill et al 2017 Neurology (12)	Prospective cohort Germany (1) 2007-2011	Spontaneous ICH maximally treated	Mortality at 3 and 12 months Unfav outcome (mRS 4-6) at 3 and 12 months	471 12-m mortality 142 (30.1%) 12-m unfav outcome 214 (45.4%)	- Age Mean 70 SD 12 - Female 45.4%	No	AUC ROC 12m mortality 0.65 12m unfav outcome 0.69	No	No	No
ICH-GS	Chuang et al. 2009 Int J Qual Health Care (3)	Retrospective cohort Taiwan (1) 2006-2007	Spontaneous ICH	Mortality at 1 month	293 Mortality 40 (14%)	- Age Mean 60.6 SD 16.7 - Female: 30%	Accuracy 78.8%	AUC ROC 0.74 (0.65-0.83)	Cut off NA (highest Youden) Se, Sp, PPV, NPV, LR+, LR- (Table 4)	HL GOF test p>0,05	No
	Naval et al. 2009 Neurol Res (8)	Registry of NICU USA (1) 1999-2006	Supra-tentorial, spontaneous ICH Excl of prior mRS 2-5	Mortality at 1 month	125 mortality 29 (23%)	- Age 63.5 range 34-90 - Female 42%	No	No	Cut off ≥ 8 (Proba 50/50) PPV: 62.8% NPV 97.6% Se 93.1% Sp 83.3%	chi square test obs/pred overestimation of mortality of 11.2% (p=0,03)	No
Max ICH score	Schmidt et al 2018 Neurology (11)	Prospective cohort Chicago US (1) 2010-2017	Spontaneous ICH	Unfav outcome (mRS 4-6) at 3 months	372 Mortality at 3m: 153 (41%) Unfav outcome at 3m: 236 (63%)	- Age Mean 67 SD 14 - Female 51%	Likelihood ratio X^2 p<0.001	AUC ROC 3m mortality 0.82 (0.78-0.86) 3m unfav outcome 0.88 (0.85-0.92)	No	- Histogram score vs observed mortality and poor outcome - Unweighted sum of squared errors test for GOF p>0.3	No

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
VASOGRADE	Dengler et al. 2017 Eur J Neurol (14)	Hosp registry, managed in ICU Germany (1) 2009-2015	aneurysmal SAH	- unfav outcome (mRS 3-6) at 12 months - unfav outcome (mRS 4-6) at 12 months	423 208 (53.1%)	- Age mean 54.2 SD 13.7 - Female 69%	No	AUC ROC mRS 3-6 0.711 mRS 4-6 0.709	No	Histogramme	No
SAH sum score / Hijdra score	Claassen et al. 2004 Crit Care Med (15)	Prospective cohort 1996-2002 USA (1)	SAH Excl AVM	Unfav outcome (mRS 4–6) at 3 months	413 Unfav outcome 40.4%	- Age mean 54 SD 14- Female 71%	No	AUC ROC0.67 (0.61-0.73)	No	No	No
SAHIT	Mascitelli et al 2018 Neurosurgey (16)	Trial cohort admitted in ICU USA (1) 2003-2007	aneurysmal SAH Excl lost to follow up (67)	- Unfav outcome (mRS 3-6) at 6 months - Mortality at 6 months	338 Mortality: 38 (10,1%) Unfav outcome: 100 (29,6%)	- Age mean 54 SD 12 - Female: NA	R^2 Brier score Brier scaled (Figure 2)	AUC ROC Unfav outcome: core: 72.8 (66.8-78.9) Neuroimaging: 73.2 (67.1-79.2) Full: 73.4 (67.5-79.4) Mortality: core: 72.1 (62.1-82.2) neuroimaging: 73.9 (64.4-83.5) Full: 74.4 (65.1-83.8)	No	calibration plot Intercept Slope (figure 2)	No
HAIR / SAH score	Witsch et al. 2016 Ann Neurol (17)	- Prospective cohort (SHOP) 1996-2014 USA (1) - CONSCIOUS-1 Israel, Europe, North America (52)	SAH Excl AVM Excl missing predictors (97) multiple Imputation for mRS (351) by MCMCM	Unfav outcome (mRS 4-6) at 12 months	- SHOP: 1526 Unfav outcome: 1200 (79%) - CONSCIOUS: 413	- SHOP: - Age mean 55.3 SD 14,5 - Female 68% -CONSCIOUS: median 55	Nagelkerke R^2 Cox/Snell R^2 - COUNSCIOUS-1: N-R^2 0.17 C/S R^2 0.11 - SHOP: N-R^2 0.45 C/S R^2 0.32	AUC ROC - CONSCIOUS-1: 71.8 (66.0-77.5) - SHOP: 88.3 (86.4-90.2)	No	No	No
	Dengler et al. 2017 Eur J Neurol (14)	Hosp registry, managed in ICU Germany (1) 2009-2015	aneurysmal SAH	- unfav outcome (mRS 3-6) at 12 months - unfav outcome (mRS 4-6) at 12 months	423 208 (53.1%)	- Age mean 54.2 SD 13.7 - Female 69%	No	AUC ROC mRS 3-6 0.739 mRS 4-6 0.737	No	Histogram	No
	Abulhasan et al. 2017 Neurocrit Care (18)	Retrospective cohort in NICU Canada (1) 2010-2016	SAH H&H 1-5 (multiple imputation MICE)	Mortality at discharge	434 Mortality 14.10%	- Age 56 48-65 - Female 63.6%	No	AUC ROC 0.89	No	Calibration curve intercept=-0.05 slope=0.77	No

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
ISAT	Dijkland et al. 2016 Crit Care Med (13)	Hosp registry, managed in ICU The Nederlands (1) 2007-2011 (same biostat. as initial publication)	Presumed aneurysmal SAH Excl lost to follow up	Mortality at 2 months	307 Mortality 94 (30.6%)	- Age Median 56 IQR 47–66 - Female 65%	No	AUC ROC WFNS at time of treatment: 0.89 WFNS at admission: 0.82	No	- Plot - intercept and slope: slopes adm WFNS 1.417 ttt WFNS 1.959 intercept WFNS adm 1.502 WFNS ttt 2.248	No
APACHE II	Claassen et al. 2004 Crit Care Med (15)	Prospective cohort 1996-2002 USA (1)	SAH Excl AVM	Unfav outcome (mRS 4–6) at 3 months	413 Unfav outcome 40.4%	- Age Mean 54 SD 14 - Female 71%	No	AUC ROC 0.66 (0.60–0.73)	No	No	No
	Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19)	Finnish Intensive Care Consortium (21) 2003 - 2012	ICH [Excl missing data on predictors or outcome (n=1479)]	Mortality at 6 months	3218 [1589 for recalib. to predict the mortality at 6m - 1629 for validation] Mortality: 1527 (48%) [Validation: 786 (48%)]	- Age: median 60 IQR 52-69 - Female: NA	No	AUC ROC 0.83 (0.81 - 0.85)	No	pvalue HL <0.001, pvalue GiViTI <0.001	Yes To predict outcome at 6 months
	Huang et al. 2012 Eur J Neurol (6)	Registry of NICU Chine (1) 2000-2011	primary pontine hemorrhage (ICH)	Mortality at 1 month	75 mortality 31 (41%)	- Age 54.8 SD 12.7 - Female 21%	No	AUC ROC 0.919 (0.843 – 0.995)	Youden Cut off 16.5 Se 91.2% Sp 86.5%	- HL GOF test p=0.428 - Table obs / expected	No
	Moon et al. 2015 J Clin Neurosci (20)	Prospective cohort adm in ICU South Korea (1) 2001-2012	ICH (60%) and ischemic strokes (40%) Excl 44 missing data	Mortality at discharge	ICH only: 300 Mortality 81 (27%)	- Age Mean 57.3 SD 17.2 - Female 48%	No	ICH only: AUC ROC 0.805	No	ICH only: - Calibration curve - HL GOF test p=0.782	No
	Rodriguez-Fernandez et al. 2018 BMJ Open (10)	Prospective multicenter Spanish Cohort (3) 2009-2012	Spontaneous ICH	Mortality at discharge	336 Mortality 181 (54%)	- Age Median 62 IQR 50-70 - Female NA	No	AUC ROC 0.80 (0.74-0.84)	No	- HL GOF test p=ns - GiViTI calibration belt p=0.43	No

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
SAPS	Handshu et al. 2005 J Neurol (22)	Handshu et al. Prospective cohort adm in ICU Germany (2) Period NA	ICH (54%) and ischemic stroke (46%) requiring endotracheal intubation	Mortality at 10 days, 3 months and 12 months	90 3-mM 58.9% 12-mM 67.8%	- Age mean 64.3 SD 10.4 - Female 50%	No	AUC ROC 10 days: 0.67 (0.55-0.80) 3months: 0.75 (0.65-0.86) 12 months: 0.77 (0.67-0.88)	Youden 10d: cut off > 12 Se 66.5% Sp 72.1% 3m: cut off > 12 Se 56.6% Sp 83.8% 12m: cut off > 10 Se 77.0% Sp 72.4%	No	No
SAPS II	Schuiling et al. 2005 Neurosurgery (23)	Retrospective cohort / The Nederlands (1) 2000 - 2002	SAH with long lasting and severe vasospasm	Unfav outcome (mRS 4-6) at 3 months	136 Unfav outcome: 65 (48%)	- Age mean 55 range 17-93 - Female 73%	No	AUC ROC 0.85 (0.78-0.91)	No	No	No
	Schuiling et al. 2005 Neurosurgery (23)	Prospective cohort adm in ICU Germany (2) Period NA	ICH (54%) and ischemic stroke (46%) requiring endotracheal intubation	Mortality at 10 days, 3 months and 12 months	90 3-mM 58.9% 12-mM 67.8%	- Age mean 64.3 SD 10.4 - Female 50%	No	AUC ROC 10 days: 0.68 (0.57 – 0.80) 3months: 0.77 (0.67 – 0.97) 12 months: 0.77 (0.66 – 0.88)	Youden 10d: cut off > 40 Se 75.9% Sp 55.7% 3m: cut off > 36 Se 84.9% Sp 62.2% 12m: cut off > 40 Se 72.1% Sp 82.8%	No	No
	Celi et al. 2012 J Pers Med (24)	Retrospective analysis on prospective cohort USA (1) 1995 - 2006	SAH	Mortality at discharge	MIMIC database 150 Mortality: 57 (25.6%)	NA	No	AUC ROC 0.84	No	HL GOF test p<0.001	No
	Czorlich et al. 2015 Acta Neurochir (Wien) (25)	ICU registry Germany (1) 2010 - 2014	aneurysmal SAH	Unfav outcome (GOS 1-3) at 1 month	263 Mortality: 49 (18.3%)	- Age mean 54,4 SD 13,74 - Female 70%	No	AUC ROC 0,834 (0,771-0,896)	No	No	No
	Moon et al. 2015 J Clin Neurosci (20)	Prospective cohort adm in ICU / South Korea (1) 2001-2012	ICH (60%) and ischemic strokes (40%) [Excl 44 missing data]	Mortality at discharge	ICH only: 300 Mortality 81 (27%)	- Age Mean 57.3 SD 17.2 - Female 48%	No	ICH only: AUC ROC 0.783	No	ICH only: - Calibration curve - HL GOF test p=0.485	No
	Huang et al. 2012 Eur J Neurol (6)	Registry of NICU Chine (1) 2000 - 2011	primary pontine hemorrhage (ICH)	Mortality at 1 month	75 mortality 31 (41%)	- Age 54.8 SD 12.7 - Female 21%	No	AUC ROC 0.890 (0.817 - 0.943)	Cut off 32.5 (Youden) Se 82.4% Sp 86.5%	- HL GOF test p=0.682 - Table obs / expected	No

Score	External validation	Study design Country (nb of centre) Date of incl.	Type of injuries	Outcomes	Sample size	Description of the population	Reporting of global performances	Reporting of discrimination	Reporting of classification	Reporting of calibration	Recalibration or updating
	Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19)	Finnish Intensive Care Consortium (21) 2003 - 2012	ICH [Excl missing data on predictors or outcome (n=1479)]	Mortality at 6 months	3218 [1589 for recalib. to predict the mortality at 6m - 1629 for validation] Mortality: 1527 (48%) [Validation: 786 (48%)]	- Age: median 60 IQR 52-69 - Female: NA	No	AUC ROC 0.84 (0.82 - 0.86)	No	pvalue HL = 0.058, pvalue GiViTI = 0.014	Yes To predict outcome at 6 months
	Barbieri et al. 2009 J Eval Clin Pract (2)	Registry of ICU Italy (1) Period NA	spontaneous intra parenchymal hemorrhage (ICH)	Mortality at 1 month	81 mortality 49 (60.5%)	- Age 64.6 SD 13.8 - Female: 36%	No	AUC ROC 0.510 (0.377 - 0.642)	No	No	No
SAPS III	Basile-Filho et al. 2018 Medicine (21)	ICU registry Brasil (1) 2011-2016	SAH	Overall mortality - Unknown horizon	51 Mortality 14 (27%)? 37.8%?	- Age Mean 54 SD 10 - Female 67%	No	AUC ROC 0.73 (0.59-0.85)	No	No	No
SOFA	Fallenius et al. 2017 Scand J Trauma Resusc Emerg Med (19)	Finnish Intensive Care Consortium (21) 2003 - 2012	ICH [Excl missing data on predictors or outcome (n=1479)]	Mortality at 6 months	3218 [1589 for recalib. to predict the mortality at 6m - 1629 for validation] Mortality: 1527 (48%) [Validation: 786 (48%)]	- Age: median 60 IQR 52-69 - Female: NA	No	AUC ROC 0.73 (0.71 - 0.76)	No	pvalue HL <0.001, pvalue GiViTI <0.001	Yes, To predict outcome at 6 months
	Basile-Filho et al. 2018 Medicine (21)	ICU registry Brasil (1) 2011-2016	SAH	Overall mortality - Unknown horizon	Mortality 14 (27%)? 37.8%?	- Age Mean 54 SD 10 - Female 67%	No	AUC ROC Day1 0.62 (0.48-0.75) Day 3 0.77 (0.63-0.87)	No	No	No
SIRS summary score	Claassen et al. 2004 Crit Care Med (15)	Prospective cohort 1996-2002 USA (1)	SAH Excl AVM	Unfav outcome (mRS 4–6) at 3 months	413 Unfav outcome 40.4%	- Age Mean 54 SD 14 - Female 71%	No	AUC ROC 0.57 (0.51–0.064)	No	No	No

References of the S4 Table

1. Edwards DF, Hollingsworth H, Zazulia AR, Diringer MN. Artificial neural networks improve the prediction of mortality in intracerebral hemorrhage. Neurology. 1999 Jul 22;53:351–7.

2. Barbieri A, Pinna C, Basso GP, Molinari R, Giuliani E, Fruggeri L, et al. Specificity and reliability of prognostic indexes in intensive care evaluation: the spontaneous cerebral haemorrhage case. J Eval Clin Pract. 2009 Apr;15:242–5.

3. Chuang YC, Chen YM, Peng SK, Peng SY. Risk stratification for predicting 30-day mortality of intracerebral hemorrhage. Int J Qual Health Care J Int Soc Qual Health Care. 2009 Dec;21:441–7.

4. Di Napoli M, Godoy DA, Campi V, del Valle M, Piñero G, Mirofsky M, et al. C-reactive protein level measurement improves mortality prediction when added to the spontaneous intracerebral hemorrhage score. Stroke. 2011 May;42(5):1230–6.

5. Godoy DA, Pinero G, Di Napoli M. Predicting mortality in spontaneous intracerebral hemorrhage: can modification to original score improve the prediction? Stroke. 2006 Apr;37:1038–44.

6. Huang KB, Ji Z, Wu YM, Wang SN, Lin ZZ, Pan SY. The prediction of 30-day mortality in patients with primary pontine hemorrhage: a scoring system comparison. Eur J Neurol. 2012 Sep;19:1245–50.

7. Maas MB, Francis BA, Sangha RS, Lizza BD, Liotta EM, Naidech AM. Refining Prognosis for Intracerebral Hemorrhage by Early Reassessment. Cerebrovasc Dis. 2017;43(3–4):110–6.

8. Naval NS, Mirski MA, Carhuapoma JR. Impact of statins on validation of ICH mortality prediction models. Neurol Res. 2009 May;31:425–9.

9. Patriota GC, Silva-Junior JM, Barcellos AC, Silva Junior JB, Toledo DO, Pinto FC, et al. Determining ICH Score: can we go beyond? Arq Neuropsiquiatr. 2009 Sep;67:605–8.

10. Rodríguez-Fernández S, Castillo-Lorente E, Guerrero-Lopez F, Rodríguez-Rubio D, Aguilar-Alonso E, Lafuente-Baraza J, 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 13;8(8):e021719.

11. Schmidt FA, Liotta EM, Prabhakaran S, Naidech AM, Maas MB. Assessment and comparison of the max-ICH score and ICH score by external validation. Neurology. 2018 04;91(10):e939–46.

12. Sembill JA, Gerner ST, Volbers B, Bobinger T, Lücking H, Kloska SP, et al. Severity assessment in maximally treated ICH patients: The max-ICH score. Neurology. 2017 Aug 1;89(5):423–31.

13. Dijkland SA, Roozenbeek B, Brouwer PA, Lingsma HF, Dippel DW, Vergouw LJ, et al. Prediction of 60-Day Case Fatality After Aneurysmal Subarachnoid Hemorrhage: External Validation of a Prediction Model. Crit Care Med. 2016 Aug;44:1523–9.

14. Dengler NF, Sommerfeld J, Diesing D, Vajkoczy P, Wolf S. Prediction of cerebral infarction and patient outcome in aneurysmal subarachnoid hemorrhage: comparison of new and established radiographic, clinical and combined scores. Eur J Neurol. 2017 Sep 23;

15. Claassen J, Vu A, Kreiter KT, Kowalski RG, Du EY, Ostapkovich N, et al. Effect of acute physiologic derangements on outcome after subarachnoid hemorrhage. Crit Care Med. 2004 Mar;32:832–8.

16. Mascitelli JR, Cole T, Yoon S, Nakaji P, Albuquerque FC, McDougall CG, et al. External Validation of the Subarachnoid Hemorrhage International Trialists (SAHIT) Predictive Model Using the Barrow-Ruptured Aneurysm Trial (BRAT) Cohort. Neurosurgery. 2018 Dec 19;

17. Witsch J, Frey H-P, Patel S, Park S, Lahiri S, Schmidt JM, et al. Prognostication of Long-Term Outcomes after Subarachnoid Hemorrhage: The FRESH Score. Ann Neurol. 2016 Jul;80(1):46–58.

18. Abulhasan YB, Alabdulraheem NS, Simoneau G, Angle M, Teitelbaum J. Mortality in spontaneous subarachnoid hemorrhage: Causality and validation of a prediction model. Neurocrit Care. 2017;27:S388.

19. Fallenius M, Skrifvars MB, Reinikainen M, Bendel S, Raj R. Common intensive care scoring systems do not outperform age and glasgow coma scale score in predicting mid-term mortality in patients with spontaneous intracerebral hemorrhage treated in the intensive care unit. Scand J Trauma Resusc Emerg Med. 2017 Oct 25;25:102.

20. Moon BH, Park SK, Jang DK, Jang KS, Kim JT, Han YM. Use of APACHE II and SAPS II to predict mortality for hemorrhagic and ischemic stroke patients. J Clin Neurosci Off J Neurosurg Soc Australas. 2015 Jan;22:111–5.

21. Basile-Filho A, Lago AF, Menegueti MG, Nicolini EA, Nunes RS, Lima SL de, et al. The use of SAPS 3, SOFA, and Glasgow Coma Scale to predict mortality in patients with subarachnoid hemorrhage: A retrospective cohort study. Medicine (Baltimore). 2018 Oct;97(41):e12769.

22. Handschu R, Haslbeck M, Hartmann A, Fellgiebel A, Kolominsky-Rabas P, Schneider D, et al. Mortality prediction in critical care for acute stroke: Severity of illness-score or coma-scale? J Neurol. 2005 Oct;252:1249–54.

23. Schuiling WJ, de Weerd AW, Dennesen PJ, Algra A, Rinkel GJ. The simplified acute physiology score to predict outcome in patients with subarachnoid hemorrhage. Neurosurgery. 2005 Aug;57:230–6; discussion 230-6.

24. Celi LA, Galvin S, Davidzon G, Lee J, Scott D, Mark R. A Database-driven Decision Support System: Customized Mortality Prediction. J Pers Med. 2012 Sep 27;2:138–48.

25. Czorlich P, Sauvigny T, Ricklefs F, Kluge S, Vettorazzi E, Regelsberger J, et al. The simplified acute physiology score II to predict hospital mortality in aneurysmal subarachnoid hemorrhage. Acta Neurochir (Wien). 2015 Dec;157:2051–9.

Table S5. Methods used to quantify the performance of the models reported in the 6 articles with development and validation studies and the 25 articles with an external validation

	Development & Validation studies N=6	**Stand-alone external validation studies N=25**
Discrimination *(ability to differentiate between patients who do or do not experience the event)*		
AUC ROC curve with 95%CI	4	16
AUC ROC curve without 95%CI	2	6
Sensitivity & specificity	1	8
Calibration *(agreement between predictions from the model and observed outcomes)*		
Hosmer–Lemeshow test	3	9
GiViTI calibration belt		2
Contingency table		1
Calibration histogram	1	4
Calibration curve & statistical tests		4
Global performance *(simultaneous evaluation of calibration and discrimination)*		
Accuracy	1	1
Brier score		1
R^2	1	3

AUC ROC: Area Under the Receiving Operative Curve; 95%CI: 95% Confident Interval