

Predicting the Neurosurgical Indication at Admission for Stroke at Bouaké University Hospital (Côte d'Ivoire): Development and Internal Validation of a Nomogram, an Integer Score and a Decision Tree

Koffi Yves Soress Dongo^{1,2*}, Faozo Stephane Landry Teti^{1,2}, Kouadio Jean-Baptiste Keke^{1,2}, Yao Bernard Fionko^{1,2}, Aderehime Haidara^{1,2}

¹Department of Neurosurgery, Bouaké University Hospital, Bouaké, Côte d'Ivoire

²Faculty of Medical Sciences, Alassane Ouattara University, Bouaké, Côte d'Ivoire

Email: *dongosoress0723@gmail.com

How to cite this paper: Dongo, K.Y.S., Teti, F.S.L., Keke, K.J.-B., Fionko, Y.B. and Haidara, A. (2026) Predicting the Neurosurgical Indication at Admission for Stroke at Bouaké University Hospital (Côte d'Ivoire): Development and Internal Validation of a Nomogram, an Integer Score and a Decision Tree. *Open Journal of Modern Neurosurgery*, 16, 401-427. <https://doi.org/10.4236/ojmn.2026.164034>

Received: September 15, 2026

Accepted: October 25, 2026

Published: October 28, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Background: In resource-limited settings, the decision to operate on a stroke rests on individual experience, with no formal tool. **Objective:** To develop and internally validate a model estimating, at admission, the probability that a patient meets a neurosurgical indication. **Patients and Methods:** Prospective cohort of 250 consecutive patients admitted for imaging-confirmed stroke between January 2022 and December 2023. The target was the operative indication recorded in the chart. A logistic regression with backward selection was translated into a nomogram and an integer score. A classification and regression tree was built in parallel. Internal validation reproduced the whole process across 1000 bootstrap resamples. **Results:** An indication was recorded in 107 patients (42.8%). Four variables were independently associated with it: hydrocephalus (adjusted odds ratio: 4.15; 95% CI: 2.03 - 8.46), mass effect (2.42), intraventricular extension (2.31) and impaired consciousness (2.23). The apparent area under the curve was 0.77 (0.71 - 0.83), the optimism-corrected area was 0.73 (95% CI: 0.66 - 0.79), and the corrected calibration slope was 0.84 (95% CI: 0.65 - 1.03). The model outperformed existing scores (ICH score: 0.66; ≤ 0.60 for the others). The probability of indication rose from 15.7% to 95.0% across the integer score (0 - 5). Of 107 indications, only 45 were operated on, and none of the 15 aneurysmal indications were operated on. **Conclusion:** A four-variable bedside tool estimates the probability of a neurosurgical indication in stroke and outperforms prognostic scores repurposed to this end. It objectifies a 58% access gap. External multicentre validation remains necessary.

Keywords

Stroke, Surgical Indication, Nomogram, Decision Tree, Predictive Model, Neurosurgery, Sub-Saharan Africa

1. Introduction

Stroke is the second leading cause of death and the third leading cause of disability worldwide. Sub-Saharan Africa bears a disproportionate share: a predominance of haemorrhagic forms, onset at a younger age, and high case fatality [1]–[10]. In this region, stroke neurosurgery faces a shortage of resources: limited technical facilities, sometimes delayed imaging, frequent absence of interventional neuroradiology, and constrained human resources [7] [9] [11] [12].

The decision to intervene—evacuation of a haematoma, decompressive craniectomy, ventricular drainage, treatment of an aneurysm—rests on a set of clinical and radiological arguments. Their integration remains largely dependent on individual experience. The randomised trials that underpin international doctrine were conducted in systems where access to surgery is not the limiting factor [13]–[19]. Their transposability to low-resource settings is uncertain. No formal tool helps the frontline practitioner to assess, at admission, the probability that a patient meets an operative indication.

Transparent, parsimonious and calibrated prediction models can structure this reasoning, harmonise practice and serve as a triage aid [20]–[22]. Two formats lend themselves to bedside use: the nomogram, which translates a regression model into a graphical points scale, and the decision tree (CART), which lays out a sequence of binary rules [21] [23]. To our knowledge, no model has been developed to predict candidacy for a neurosurgical intervention in stroke in sub-Saharan Africa: the existing tools (ICH score, NIHSS, WFNS, Fisher) target prognosis, not the operative decision. The latter follows its own logic, that of correctable mechanical reperfusion, which a prognostic score captures poorly.

We developed and internally validated a tool estimating the probability that a patient is a candidate for an intervention, from a prospective cohort of 250 patients managed in the department of neurosurgery of Bouaké University Hospital (Côte d’Ivoire). The variables used, clinical and radiological, are all available at admission. We propose two versions of it—a nomogram with its derived integer score, and a decision tree—and we report its discrimination, calibration, clinical benefit and robustness, comparing them with the existing prognostic scores. The model also serves to quantify unmet neurosurgical need. Reporting follows the STROBE and TRIPOD recommendations [20] [24].

2. Methods

2.1. Study Design and Setting

This was a predictive-model development analysis, carried out on a single-centre

prospective cohort assembled in the department of neurosurgery of Bouaké University Hospital (CHU), Côte d'Ivoire, the reference facility for the centre of the country. The methodology of cohort assembly, the collection procedures and the definitions have been detailed in Sections 2.2 to 2.4. The present work is reported in accordance with the STROBE [24] and TRIPOD (development and internal validation of a prediction model) [20] recommendations.

2.2. Participants

The source population comprised consecutive patients admitted to the Department of Neurosurgery of Bouaké University Hospital, the referral neurosurgical facility for the centre of the country, for imaging-confirmed stroke (computed tomography or magnetic resonance imaging), from 1 January 2022 to 31 December 2023. Patients aged 18 years or over were eligible; traumatic or tumour-related haemorrhages and patients without imaging were excluded. Of the 250 patients, 93 (37.2%) had been referred by another health facility and 53 (21.2%) arrived by medicalised transport; the median delay between onset and admission was 1 day (IQR: 1 - 3; available in 248 patients). The analysis covered the 250 patients of the cohort: 149 intracerebral haemorrhages (59.6%), 65 ischaemic strokes (26.0%), 30 subarachnoid haemorrhages (12.0%) and 6 strokes of another mechanism (2.4%, cerebral venous thrombosis and other causes), *i.e.*, $149 + 65 + 30 + 6 = 250$; all subgroup analyses refer to these counts. None had a missing value on the target or on the four predictors retained in the final model.

Sample size was not calculated a priori: it corresponds to all consecutive patients of the parent cohort over the study period; its adequacy for model development was assessed a posteriori by the number of events per variable (Section 2.5).

2.3. Target: The Neurosurgical Indication

The target to be predicted was the operative indication, as recorded in the chart by the neurosurgical team at the end of the admission work-up (binary variable), on the basis of the available clinical and imaging data. The indication was initially assessed by the on-call neurosurgeon, then discussed collegially with a second neurosurgeon when the situation was complex or when several therapeutic options were conceivable. This is an expertise-based target, reflecting actual practice in a resource-limited setting, not a universal gold standard. The model predicts the probability that a patient meets the indication criteria applied in our centre; it does not prejudge the superiority of surgery over medical treatment. This choice was preferred over surgery actually performed, which also depends on the availability of the technical facility—a determinant of access to care, distinct from the medical decision itself. The tool aims to estimate the probability of a neurosurgical indication independently of this access constraint. Surgery actually performed was nevertheless analysed as an alternative target in a sensitivity analysis. Model development is summarised in **Figure 1**.

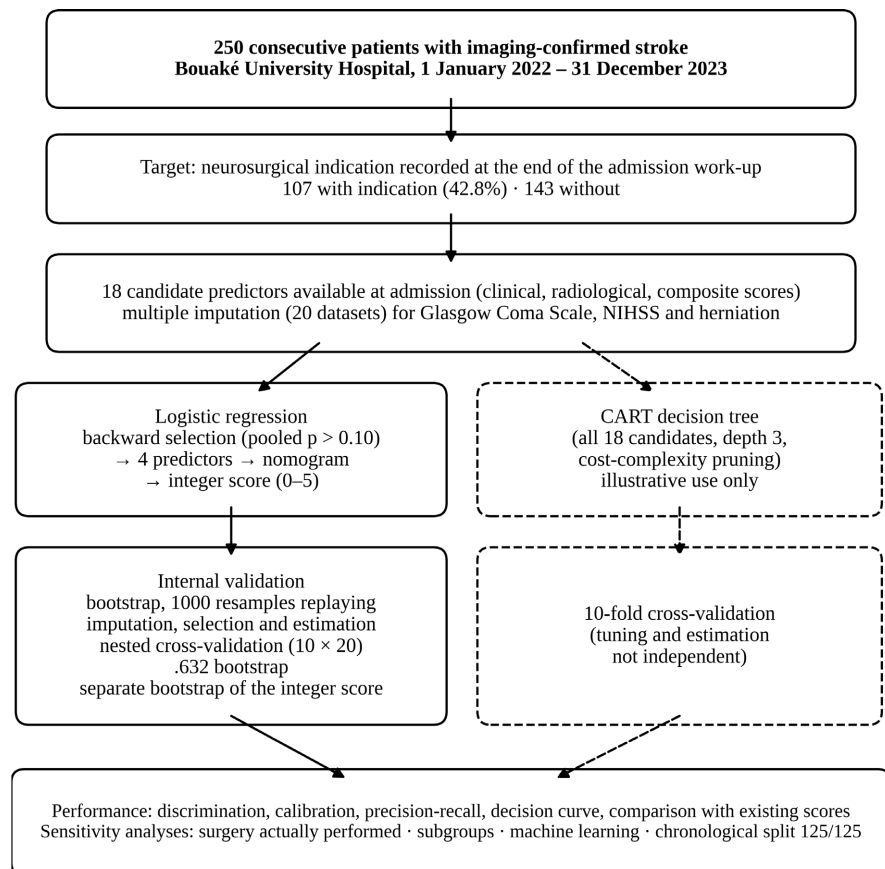


Figure 1. Model development and validation flow diagram (TRIPOD schema). The 250 patients, with no missing data on the target or on the final predictors, were used to develop the nomogram (4 variables) and the CART tree, then for internal validation.

2.4. Candidate Predictors

The predictors, all available at admission, were chosen a priori on pathophysiological and feasibility grounds. For the clinical domain: age, sex, Glasgow Coma Scale, impaired consciousness, motor deficit, anisocoria, vomiting. For the radiological domain: stroke type, intracerebral haemorrhage, haematoma volume, mass effect, herniation, hydrocephalus, intraventricular extension. For the composite scores: ICH score, WFNS, modified Fisher, NIHSS. Haematoma location (missing in 148 of the 149 haemorrhages) and lesion territory (not recorded in 249 patients) were excluded at the outset. The variables defined only for haemorrhages—volume, herniation, ICH score, WFNS, modified Fisher—were coded zero where they do not apply, without a domain indicator: this coding conflates absence of lesion and absence of measurement (see Limitations).

2.5. Model Development

Nomogram. A logistic regression model was developed by stepwise backward selection (removal at $p > 0.10$). Candidate variables with missing data (Glasgow Coma Scale 12.4%; NIHSS 1.6%; herniation 3.6%) were handled by multiple imputation: twenty datasets were generated from the whole set of candidate predic-

tors and the target, and selection was conducted on the p -values pooled according to Rubin's rules, a variable being removed when this value exceeded 0.10; none of the variables concerned was ultimately retained: the final model was therefore estimated on the 250 complete observations, without imputation. With 107 events for four predictors, the number of events per variable was 26.8 in the final model, but only 5.9 relative to the 18 candidate predictors listed above; it is at this level that the optimism forms which internal validation was designed to capture. As the four predictors were binary, the linearity-of-logit assumption did not apply. Collinearity was checked by the variance inflation factor. The final model was translated into a nomogram, then into a simplified integer score.

Decision tree. A classification and regression tree (CART) was built in parallel, on the whole set of 18 candidate predictors listed in Section 2.4—clinical and radiological variables and the composite scores (ICH score, WFNS, modified Fisher, NIHSS), the latter coded zero outside their domain of definition—to expose possible thresholds and interactions. The ICH score could therefore enter the tree although it was not retained in the nomogram; as it is coded zero outside intracerebral haemorrhage, a split on this score partly separates haemorrhages from the other mechanisms (see Limitations). Pruning was performed by cost-complexity, the penalisation parameter being chosen by 10-fold cross-validation maximising the AUC. A deliberately shallow tree (depth 3) was retained for clinical presentation.

2.6. Internal Validation, Performance and Clinical Benefit

Discrimination was measured by the area under the ROC curve (AUC; DeLong interval), complemented by the precision-recall curve. Calibration was assessed by the Hosmer-Lemeshow test, the Brier score, the calibration slope and intercept, and a smoothed curve with a bootstrap interval. Internal validation reproduced the entire development process—imputation, backward selection and coefficient estimation—in each of 1000 bootstrap resamples, optimism being defined as the mean difference between performance in the bootstrap sample and performance in the original sample; a nested repeated cross-validation (10 folds $\times$ 20 repetitions) and a 0.632 bootstrap completed the analysis. Clinical benefit was assessed by decision curve analysis, and discrimination was compared with that of existing scores repurposed for predicting the indication. For the tree, performance was estimated by 10-fold cross-validation; as the choice of the complexity parameter and performance estimation were not conducted as two independent procedures, this value constitutes a cross-validation estimate and not a performance on an independent sample. The integer score was validated in a dedicated bootstrap of 1000 resamples: in each resample, the coefficients of the four predictors were re-estimated, rounded into integer weights by the same rule as in the original sample (division by the smallest coefficient, then rounding to the nearest integer), the correspondence between score and probability was re-derived by logistic regression on the score, and the Youden-optimal threshold was re-selected; the resulting score was

then applied to the original sample. The optimism of the area under the curve, of the calibration slope and intercept, of the Brier score and of sensitivity and specificity at the selected threshold was estimated as for the nomogram. As this procedure did not repeat variable selection, the continuous four-variable model was submitted to the same procedure, so as to isolate the share of optimism due to rounding and threshold choice.

2.7. Statistical Analysis

Continuous variables are described by their median (interquartile range) and compared by the Mann-Whitney test. Categorical variables are described by counts and percentages, compared by the χ^2 test. Associations are expressed as crude and adjusted odds ratios (OR), with 95% confidence intervals. The significance threshold was set at 0.05 (two-sided). Analyses were performed with Python (pandas, statsmodels, scikit-learn).

2.8. Ethical Considerations

Bouaké University Hospital did not have, over the study period, a constituted ethics committee. The study was conducted after authorisation from the head of the department of neurosurgery and the hospital management, in accordance with the principles of the Declaration of Helsinki, and has since received the approval of the Institutional Ethics Committee of Bouaké University Hospital (approval no. CEI/CHU-BKE/102/2026). The data were analysed anonymously; individual consent was not required, as the data had been collected as part of routine care, this waiver being covered by the approval obtained.

3. Results

3.1. Population and Frequency of the Surgical Indication

A neurosurgical indication was retained in 107 of the 250 patients (42.8%). These patients had the same age (median 56 years) and the gender-related distribution (62% men) as the others, but showed more marked signs of severity: a lower Glasgow Coma Scale (11 vs 12; $p = 0.035$), a higher ICH score (2 vs 0; $p < 0.001$), a larger haematoma (32 vs 3 mL; $p < 0.001$), a greater cerebral herniation (median 4 vs 0 mm; $p < 0.001$), a higher NIHSS (19 vs 17; $p = 0.011$) and more frequent anisocoria (24.3% vs 13.3%; $p = 0.038$). Stroke mechanism differed between the two groups ($p < 0.001$): patients with an indication comprised 77 intracerebral haemorrhages (72.0%), 14 ischaemic strokes (13.1%), 15 subarachnoid haemorrhages (14.0%) and 1 stroke of another mechanism (0.9%), versus 72 (50.3%), 51 (35.7%), 15 (10.5%) and 5 (3.5%) in the other patients (**Table 1**). Read by mechanism, an indication was retained in 77 of the 149 intracerebral haemorrhages (51.7%), 14 of the 65 ischaemic strokes (21.5%), 15 of the 30 subarachnoid haemorrhages (50.0%) and 1 of the 6 other strokes. Hydrocephalus (41.1% vs 11.2%), intraventricular extension (55.1% vs 23.8%), mass effect (67.3% vs 41.3%) and impaired consciousness (63.6% vs 36.4%) were all more frequent in patients with an

indication ($p < 0.001$ for each).

Table 1. Patient characteristics according to the presence of a neurosurgical indication.

Characteristic	Indication (n = 107)	No indication (n = 143)	<i>p</i>
Age, years, median (IQR)	56 (45 - 63)	56 (43 - 66)	0.98
Male sex, n (%)	66 (61.7)	89 (62.2)	1.00
Glasgow Coma Scale, median (IQR)	11 (8 - 13)	12 (10 - 14)	0.035
Impaired consciousness, n (%)	68 (63.6)	52 (36.4)	<0.001
Motor deficit, n (%)	86 (80.4)	100 (69.9)	0.084
Anisocoria, n (%)	26 (24.3)	19 (13.3)	0.038
Vomiting, n (%)	35 (32.7)	40 (28.0)	0.50
NIHSS, median (IQR)	19 (15 - 23)	17 (12 - 21)	0.011
Stroke mechanism, n (%)			<0.001
Intracerebral haemorrhage	77 (72.0)	72 (50.3)	
Ischaemic stroke	14 (13.1)	51 (35.7)	
Subarachnoid haemorrhage	15 (14.0)	15 (10.5)	
Other mechanism	1 (0.9)	5 (3.5)	
Haematoma volume, mL, median (IQR)	32 (0 - 49)	3 (0 - 40)	<0.001
Herniation, mm, median (IQR)	4 (0 - 7)	0 (0 - 5)	<0.001
Mass effect, n (%)	72 (67.3)	59 (41.3)	<0.001
Hydrocephalus, n (%)	44 (41.1)	16 (11.2)	<0.001
Intraventricular extension, n (%)	59 (55.1)	34 (23.8)	<0.001
ICH score, median (IQR)	2 (0 - 3)	0 (0 - 2)	<0.001

Column percentages. Stroke mechanism: 149 + 65 + 30 + 6 = 250 patients. Haematoma volume, herniation and ICH score are coded zero outside intracerebral haemorrhage (Section 2.4); their medians refer to the whole cohort. IQR: interquartile range. *p*: Mann-Whitney test for continuous variables, χ^2 test for categorical variables.

3.2. Univariate Determinants

In univariate analysis, hydrocephalus (OR: 5.54; 95% CI: 2.90 - 10.59), intraventricular extension (3.94; 2.29 - 6.77), impaired consciousness (3.05; 1.81 - 5.14), mass effect (2.93; 1.74 - 4.94) and intracerebral haemorrhage (2.53; 1.48 - 4.32) were strongly and significantly associated with the indication. Haematoma volume (OR: 1.02 per mL; 1.01 - 1.03), herniation (1.16 per mm; 1.07 - 1.25), ICH score (1.65 per point; 1.32 - 2.05), NIHSS (1.06 per point; 1.01 - 1.10), Glasgow Coma Scale (0.90 per point; 0.82 - 0.98) and anisocoria (2.09; 1.09 - 4.03) were also significantly associated, with more modest magnitudes. Ischaemic stroke was associated with a lower frequency of indication (OR: 0.27; 0.14 - 0.52). Age, sex, motor deficit (OR: 1.76; 0.97 - 3.20), vomiting and subarachnoid haemorrhage (OR: 1.39; 0.65 - 2.99) did not reach the significance threshold (**Table 2**).

Table 2. Univariate analysis of the candidate predictors.

Variable	Crude OR	95% CI	<i>p</i>
Hydrocephalus	5.54	2.90 - 10.59	<0.001
Intraventricular extension	3.94	2.29 - 6.77	<0.001
Impaired consciousness	3.05	1.81 - 5.14	<0.001
Mass effect	2.93	1.74 - 4.94	<0.001
Intracerebral haemorrhage*	2.53	1.48 - 4.32	<0.001
Ischaemic stroke*	0.27	0.14 - 0.52	<0.001
Subarachnoid haemorrhage*	1.39	0.65 - 2.99	0.40
Haematoma volume (per mL)	1.02	1.01 - 1.03	<0.001
Herniation (per mm)	1.16	1.07 - 1.25	<0.001
ICH score (per point)	1.65	1.32 - 2.05	<0.001
NIHSS (per point)	1.06	1.01 - 1.10	0.009
Glasgow Coma Scale (per point)	0.90	0.82 - 0.98	0.018
Age (per year)	1.00	0.98 - 1.02	0.91
Male sex	0.98	0.58 - 1.64	0.93
Motor deficit	1.76	0.97 - 3.20	0.063
Anisocoria	2.09	1.09 - 4.03	0.027
Vomiting	1.25	0.73 - 2.16	0.42

OR: crude odds ratio from univariable logistic regression; CI: Wald confidence interval.

*Each stroke mechanism is compared with all other mechanisms combined (reference = the remaining patients of the cohort).

3.3. Multivariable Model and Nomogram

After this selection, four variables kept an effect of their own on the indication, independently of one another: hydrocephalus (aOR: 4.15; 95% CI: 2.03 - 8.46; $p < 0.001$), mass effect (2.42; 1.35 - 4.36; $p = 0.003$), intraventricular extension (2.31; 1.26 - 4.23; $p = 0.007$) and impaired consciousness (2.23; 1.25 - 3.96; $p = 0.007$). The Glasgow Coma Scale, the haematoma volume and the composite scores provided no further information once these four markers were taken into account; they were therefore discarded from the final model. The four variables retained were not redundant with one another either (variance inflation factors all ≤ 1.8) (**Table 3**; details in **Table 4**). The model was then translated into a nomogram (**Figure 2**), where each variable receives a number of points proportional to its weight: 100 points for hydrocephalus, 62 for mass effect, 59 for intraventricular extension and 56 for impaired consciousness.

The model equation is as follows, where *H*, *M*, *V* and *C* denote hydrocephalus, mass effect, intraventricular extension and impaired consciousness, each coded 0 (absent) or 1 (present):

$$\text{logit}(p) = -1.830 + 1.422H + 0.886M + 0.837V + 0.800C$$

$$p = \frac{1}{1 + e^{-\text{logit}(p)}}$$

The equivalent integer score assigns 2 points to hydrocephalus and 1 point to each of the three other variables (total 0 to 5).

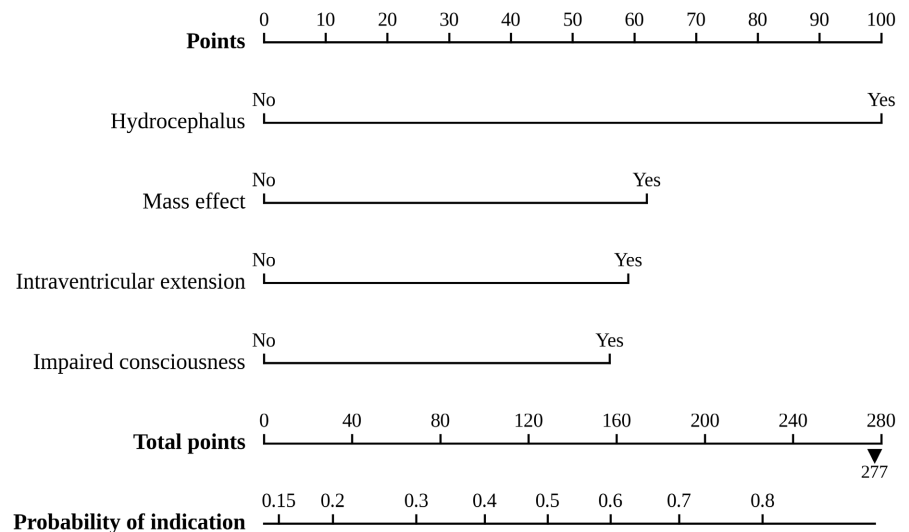


Figure 2. Nomogram for estimating the probability of a neurosurgical indication. The points of each variable are summed on the “Total points” rule (maximum attainable 277); the lower rule converts this total into a probability according to $\text{logit}(p) = -1.830 + 0.01422 \times \text{total points}$.

Table 3. Final multivariable model and nomogram points.

Variable	Adjusted OR	95% CI	<i>p</i>	Points (nomogram)
Hydrocephalus	4.15	2.03 - 8.46	<0.001	100
Mass effect	2.42	1.35 - 4.36	0.003	62
Intraventricular extension	2.31	1.26 - 4.23	0.007	59
Impaired consciousness	2.23	1.25 - 3.96	0.007	56

Table 4. Parameters of the logistic model and collinearity.

Variable	Coefficient β	Standard error	VIF
Constant	-1.830	—	—
Hydrocephalus	1.422	0.364	1.47
Mass effect	0.886	0.299	1.78
Intraventricular extension	0.837	0.309	1.80
Impaired consciousness	0.800	0.294	1.79

Standard errors derived from the 95% confidence intervals of the adjusted odds ratios. VIF: variance inflation factor.

3.4. Simplified Integer Score

For simpler bedside use, the model coefficients were rounded into an integer score: 2 points for hydrocephalus, 1 point each for mass effect, intraventricular extension and impaired consciousness (total 0 to 5). The higher the score, the more steadily

the observed probability of indication increases, from 15.7% at 0 points up to 95.0% at 5 points, in agreement with the predicted probabilities. The only exception: among the 18 patients with a score of 4, the observed rate of surgical indication (66.7%) was eleven points below the rate predicted by the model for that score (Table 5, Figure 3). The integer score was validated in its own right. Its area under the ROC curve, 0.77, identical to that of the nomogram, is the reference value of the score. In the dedicated bootstrap, the optimism of this area was 0.009, identical to that of the continuous four-variable model submitted to the same procedure (0.010): rounding the coefficients and choosing the threshold generated no measurable additional optimism. As this procedure did not repeat variable selection, the optimism attributable to selection, estimated for the nomogram (0.041), must be added: after correction for the whole development process, the area of the integer score would be about 0.73, as for the nomogram. Calibration of the score corrected for rounding and threshold choice was satisfactory (slope 0.95, intercept 0.00, Brier score 0.20); once selection is accounted for, the corrected slope of the nomogram (0.84) remains the more prudent reference. At the threshold re-selected in each resample, corrected sensitivity and specificity were 75.5% and 64.3% (apparent values 77.6% and 65.7%). The optimal threshold was, however, unstable across resamples (score ≥ 3 in 37% of them, ≥ 2 in 26%, ≥ 4 in 19%), and the rounded weighting varied, hydrocephalus receiving a weight at least equal to that of every other variable in 89% of resamples. The thresholds of the score reported in Section 3.8 are therefore operating points to be chosen according to use, not a validated cut-off. As the score is a monotone grouping of the linear predictor, its discrimination cannot exceed that of the continuous model; the equality of their apparent areas is an empirical result.

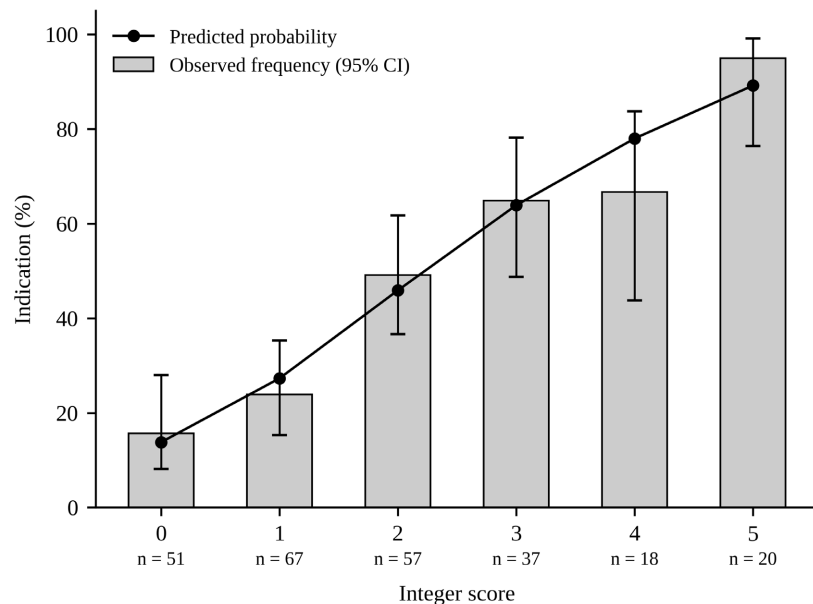


Figure 3. Risk gradient by integer score (0 - 5). Bars: observed frequency of indication with its 95% Wilson confidence interval; line: probability predicted by the model; the number of patients at each score level is given under the axis.

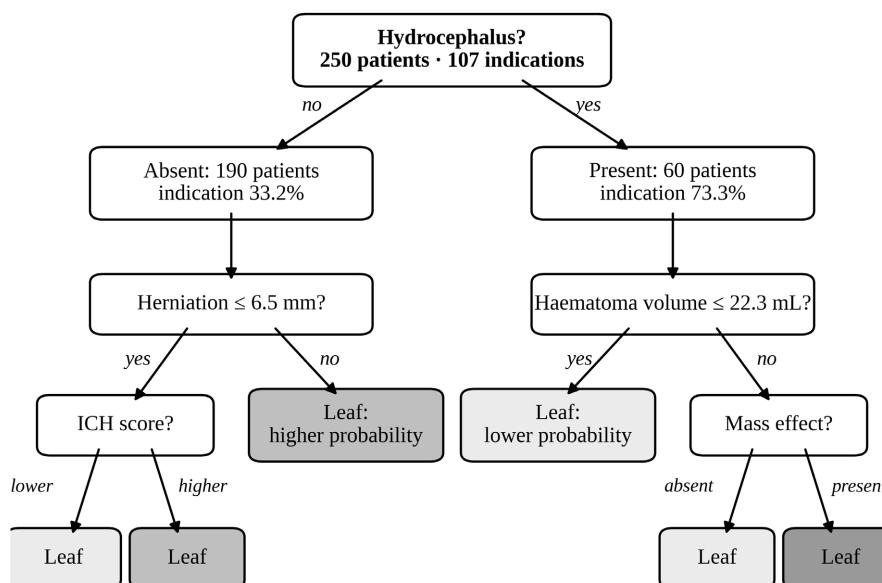
Table 5. Correspondence between the integer score, the observed frequency of indication and the probability predicted by the model.

Score	Patients (n)	Indications (n)	Observed frequency (%)	Predicted probability (%)
0	51	8	15.7	13.8
1	67	16	23.9	27.3
2	57	28	49.1	45.9
3	37	24	64.9	63.9
4	18	12	66.7	78.0
5	20	19	95.0	89.2
Total	250	107	42.8	—

The predicted probability is the count-weighted mean of the probabilities of the profiles grouped at each score level.

3.5. Decision Tree (CART)

In the classification tree, the first question asked is the presence or absence of hydrocephalus. In its absence, the decision rests on cerebral herniation and the ICH score; in its presence, it is the haematoma volume and the mass effect that guide it. The thresholds retained are clinically plausible (volume ≈ 22 mL, herniation ≈ 6.5 mm) (Figure 4). Its performance was good on the training data (AUC: 0.78), but markedly poorer in cross-validation (0.69). This gap reflects the overfitting inherent to single decision trees, which sometimes fit too precisely the data used to build them. This is why the nomogram, which is more stable, is proposed as the first-line tool, the tree serving only as a teaching aid.



Leaf shading: darker = higher probability of indication. Cross-validation AUC = 0.69.

Figure 4. CART decision tree (depth 3). Hydrocephalus is the first node; in its absence, herniation and the ICH score guide the decision; in its presence, volume and mass effect. Cross-validation AUC = 0.69.

3.6. Discrimination, Calibration and Internal Validation

The nomogram showed an apparent AUC of 0.77 (95% CI: 0.71 - 0.83) (**Figure 5**). Internal validation, which reproduced the entire development, gave a mean optimism of 0.041 and an AUC corrected to 0.73 (95% CI: 0.66 - 0.79); nested cross-validation retrieved 0.74, as did the 0.632 bootstrap. The four predictors of the final model were selected in a large majority of the resamples (hydrocephalus 96.8%; mass effect 88.7%; impaired consciousness 86.1%; intraventricular extension 84.3%), which attests to the stability of the selection. Discrimination is therefore moderate, which remains appreciable for a four-binary-variable tool that can be computed from memory. Calibration was satisfactory: corrected slope 0.84 (95% CI: 0.65 - 1.03), corrected intercept -0.01 —a slope below 1 reflecting moderate overfitting, but whose confidence interval includes 1, so that perfect calibration is not excluded. The Hosmer-Lemeshow test was not significant ($p = 0.92$), but its interpretation remains limited, with the sixteen profiles predicted by the model being very unevenly distributed over the cohort. The observed/expected ratio was 1.00, which is expected by construction. The Brier score was 0.19, a 22.4% improvement over the null model. The observed-predicted concordance was satisfactory, except at score 4 (**Figure 6** and **Figure 7**). The precision-recall curve (mean precision 0.73; prevalence 0.43) confirmed the decisional relevance of the model (**Figure 8**). The Youden index of the nomogram was maximal for any cut-off above 0.280 and up to 0.399, the predicted probability of isolated hydrocephalus; within this interval, the nomogram classifies patients exactly as an integer score ≥ 2 does (sensitivity 77.6%, specificity 65.7%), since every profile scoring 2 points or more has a predicted

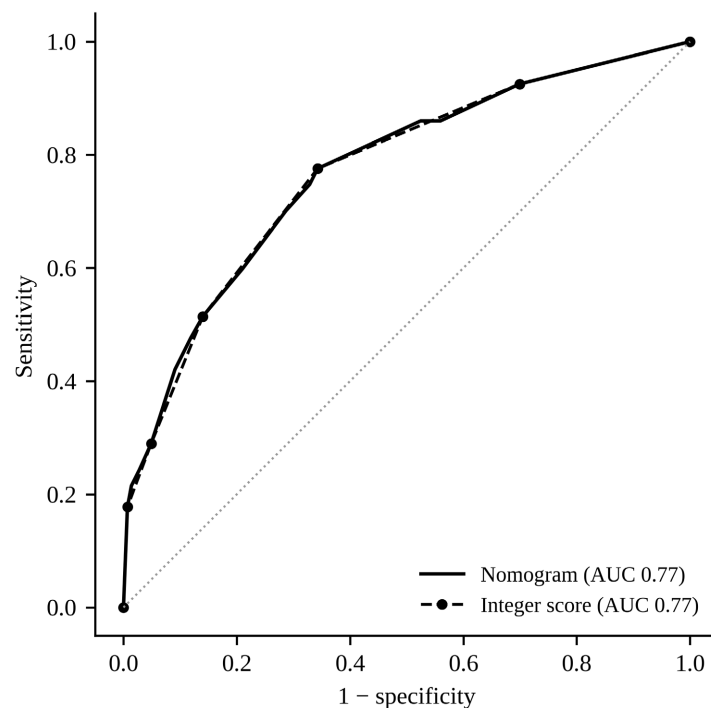


Figure 5. ROC curves of the nomogram and of the integer score.

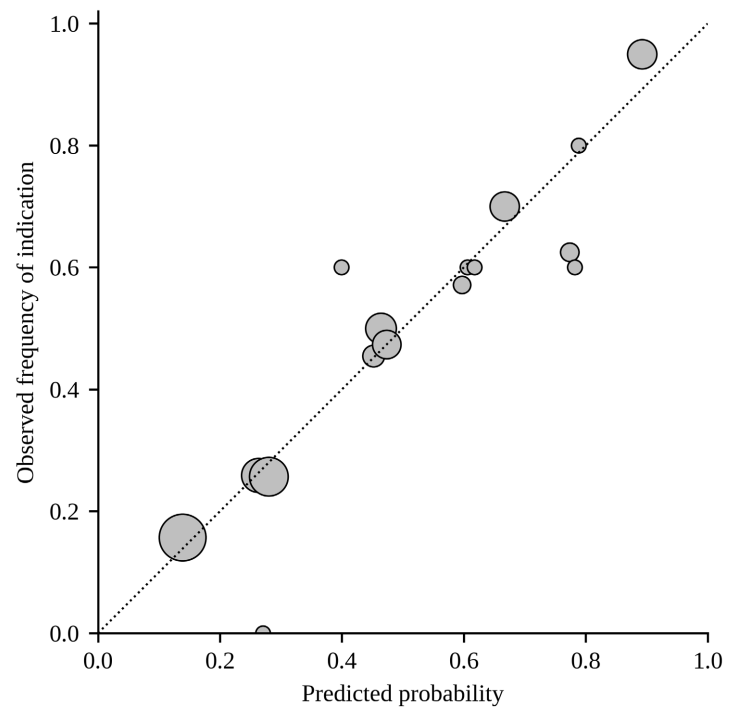


Figure 6. Calibration plot of the nomogram: observed frequency of indication as a function of the predicted probability, for each of the sixteen profiles of the model, the area of the symbols being proportional to the count (Hosmer-Lemeshow $p = 0.92$).

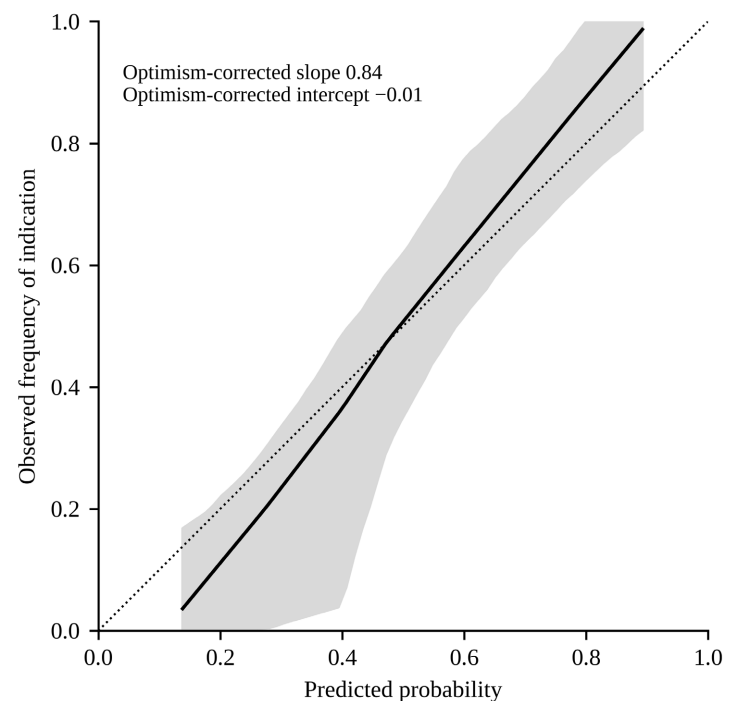


Figure 7. Internal calibration of the logistic model. The solid line represents the smoothed relationship between predicted probabilities and observed frequency; the shaded area represents the 95% confidence interval obtained by bootstrap. The optimism-corrected calibration slope was 0.84 and the corrected intercept -0.01 . The whole variable-selection process was repeated in each resample.

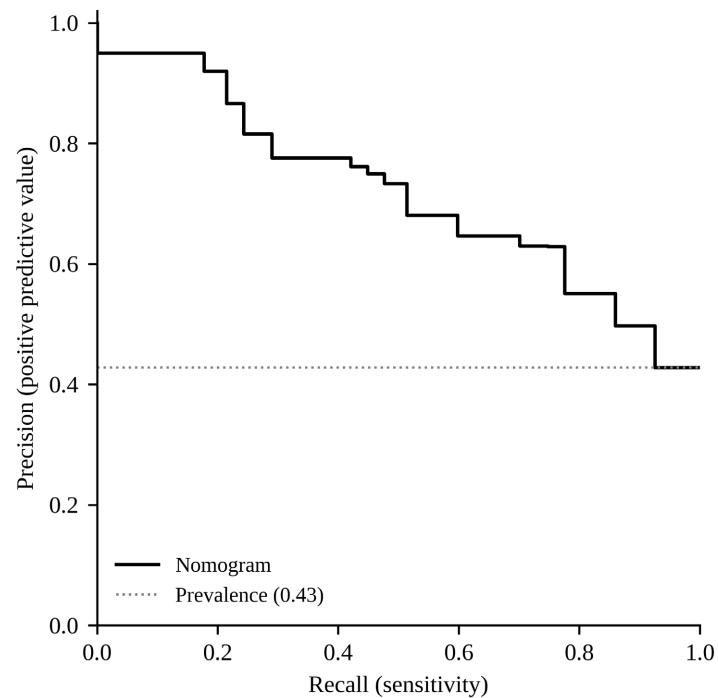


Figure 8. Precision-recall curve of the nomogram (mean precision: 0.73; prevalence: 0.43).

probability of at least 0.399 and every profile scoring 0 or 1 point has a probability of at most 0.280. Rounded to 0.40, the cut-off would exclude isolated hydrocephalus (39.9%), which scores 2 points; the operating threshold is therefore expressed on the scale of the integer score.

3.7. Clinical Benefit and Comparison with Existing Scores

On decision curve analysis, the nomogram provided a net benefit superior to the “operate on all” and “operate on none” strategies over the range of probabilities that the model predicts (0.14 to 0.89) (**Figure 9**); it was also superior to a model based on the ICH score alone. As the reference standard is the indication set by the team, this benefit measures the usefulness of reproducing that decision, not a demonstrated gain in patient outcome. The nomogram outperformed the prognostic scores repurposed for this prediction: ICH score 0.66, NIHSS 0.60, Glasgow 0.58, WFNS 0.52, modified Fisher 0.52, versus 0.77 for the nomogram. The differences all went in the same direction on the DeLong test ($p \leq 0.001$) (**Table 6, Figure 10**); this test, applied to the apparent area of the nomogram and ignoring the uncertainty attached to variable selection, is anti-conservative here. This comparison is indicative: it pits the apparent area of the nomogram against scores none of whose parameters were estimated on our cohort, and three of which are defined only for a subset of patients—their coding at zero outside their domain of definition mechanically lowers their discrimination over the whole cohort; relative to the corrected AUC (0.73), the hierarchy nonetheless remains unchanged.

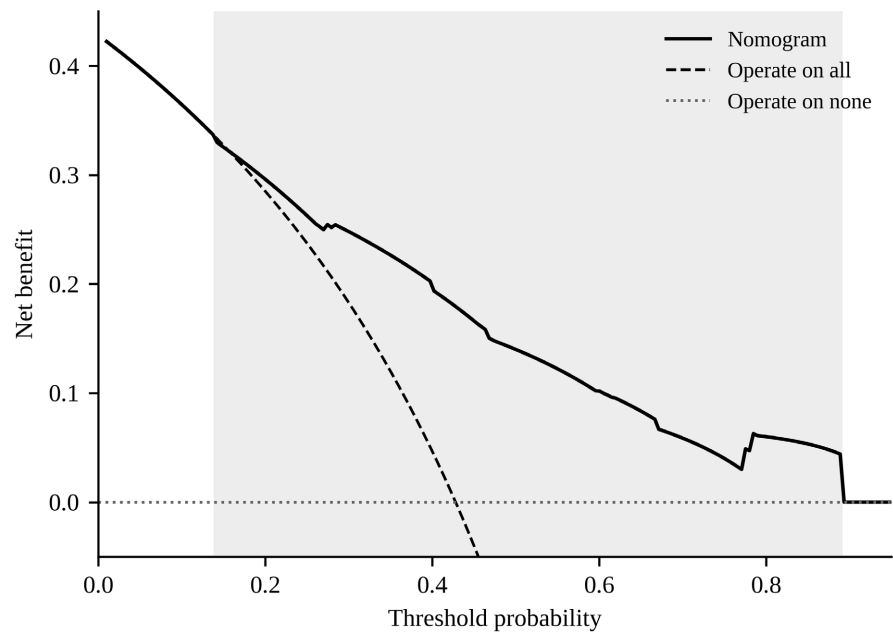


Figure 9. Decision curve analysis. Net benefit of the nomogram (solid line) compared with the “operate on all” (dashed line) and “operate on none” (dotted line) strategies. The shaded band marks the range of probabilities predicted by the model (0.14 to 0.89), over which the nomogram provides the higher net benefit.

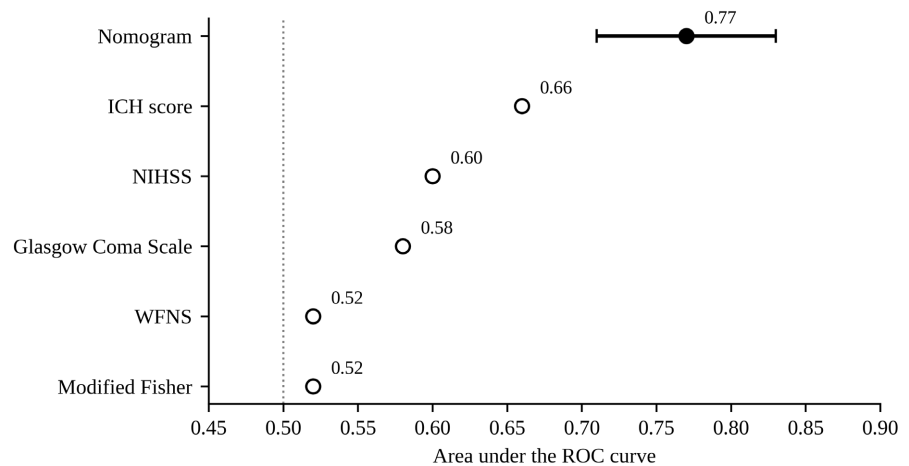


Figure 10. Areas under the ROC curve for predicting the surgical indication: nomogram (0.77, with its 95% confidence interval, 0.71 - 0.83) and existing scores repurposed for this prediction (ICH score, NIHSS, Glasgow Coma Scale, WFNS, modified Fisher). The dotted line at 0.50 corresponds to absence of discrimination.

Table 6. Discrimination of the nomogram and of existing scores for predicting the neurosurgical indication.

Tool	Original purpose	Area under the ROC curve	Comparison with the nomogram
Nomogram	neurosurgical indication	0.77 (95% CI: 0.71 - 0.83)	—

Continued

ICH score	vital prognosis of intracerebral haemorrhage	0.66	$p \leq 0.001$
NIHSS	severity of the neurological deficit	0.60	$p \leq 0.001$
Glasgow Coma Scale	depth of impaired consciousness	0.58	$p \leq 0.001$
WFNS	severity of subarachnoid haemorrhage	0.52	$p \leq 0.001$
Modified Fisher	subarachnoid haemorrhage burden	0.52	$p \leq 0.001$

DeLong test. Indicative comparison: the nomogram area is apparent, and three of the comparators are defined only for a subset of the cohort.

3.8. Operational Thresholds of the Score

The thresholds of the integer score offer operating points suited to different uses (**Table 7**); these are apparent values, the optimism-corrected performance at the re-selected threshold being given in Section 3.4. A low threshold (≥ 1) favours sensitivity (92.5%), useful so as not to miss a candidate to refer. An intermediate threshold (≥ 2) balances sensitivity (77.6%) and specificity (65.7%). A high threshold (≥ 5) reaches 99.3% specificity and 95.0% positive predictive value (likelihood ratio 25.4).

Table 7. Operational thresholds of the integer score.

Threshold	Sensitivity	Specificity	PPV	LR+
≥ 1	92.5%	30.1%	49.7%	1.3
≥ 2	77.6%	65.7%	62.9%	2.3
≥ 5	17.8%	99.3%	95.0%	25.4

3.9. Unmet Indications

Of the 107 indications recorded, only 45 (42%) led to an intervention. There remained, therefore, 62 unmet indications. The gap varied strongly according to mechanism (**Table 8**): total for subarachnoid haemorrhage (none of the 15 aneurysmal indications treated, neither by clipping nor by embolisation, for want of interventional neuroradiology), 64% for malignant infarction (5 craniectomies out of 14), 48% for intracerebral haemorrhage (40 operated on out of 77), and the single indication belonging to another mechanism also remained unanswered. These 45 patients received 54 procedures, some having had combined procedures: 24 external ventricular drains, 24 decompressive craniectomies and 6 evacuations. The numbers of procedures therefore do not add up like counts of distinct patients.

Table 8. Unmet indications by mechanism.

Mechanism	Indications recorded	Operated on	Gap
Subarachnoid haemorrhage	15	0	100%
Malignant infarction	14	5	64%

Continued

Intracerebral haemorrhage	77	40	48%
Other	1	0	100%
Total	107	45	58%

3.10. Sensitivity Analysis: Surgery Actually Performed

When the target was surgery actually performed (45 patients operated on), the same severity markers remained decisive. The relative weight nevertheless shifted towards mass effect (aOR: 5.00; 2.04 - 12.29) and impaired consciousness (aOR: 5.85; 2.41 - 14.24). Hydrocephalus remained significant (aOR: 2.42; 1.08 - 5.44), whereas intraventricular extension lost its significance (aOR: 1.41; 0.65 - 3.07). This model discriminated even better (AUC: 0.83), which is expected since operated patients represent the most severe end of the spectrum. The shift of relative weight towards mass effect and impaired consciousness, and the loss of significance of intraventricular extension, indicate that the target “surgery actually performed” selects a more severe profile: these determinants do not exactly overlap with those of the indication and should not be read as a mere confirmation of the main model.

3.11. Subgroup Performance

Discrimination was homogeneous by sex (AUC: 0.75 in men, 0.79 in women) and by age (0.78 before 56 years, 0.75 beyond). It was highest in intracerebral haemorrhage (AUC: 0.77; $n = 149$), the subgroup that concentrates most indications. It was more modest in ischaemia (0.62; $n = 65$) and subarachnoid haemorrhage (0.62; $n = 30$), where the small number of events and the specificity of the decisional mechanisms (craniectomy for malignant infarction, aneurysmal treatment) limit the reach of the model. These subgroup performances, based on small numbers, remain exploratory.

3.12. Comparison with Machine Learning and Temporal Validation

Two complementary checks support the choice of a simple model. First, on the same four variables, two machine-learning methods did no better than logistic regression in repeated cross-validation: random forest AUC: 0.73, gradient boosting: 0.74, versus 0.76 for logistic regression fitted on those same four variables, without new selection. As these methods addressed only the four already-retained variables—sixteen possible profiles—this result mainly shows that no residual interaction is exploitable among them, rather than the general uselessness of complex approaches upstream of selection. Second, an exploratory temporal validation was carried out: the model was developed on the first chronological half of the cohort (125 patients) and evaluated on the second (125 patients). It retrieved an AUC of 0.82 on this later and independent period, despite a higher frequency of indication in the second period. This result remains an internal validation—the 125/125 split separates inclusions in chronological order and does not correspond to a validation by calendar

year—and cannot be interpreted as a demonstration of temporal robustness.

3.13. Analysis of Discordant Cases

At the score ≥ 2 threshold, the 24 false negatives (indication recorded despite a low score) had, by construction, no hydrocephalus and few signs of repercussion on CT. A third were infarctions (8/24), whose indication—decompressive craniectomy for malignant infarction—rests on clinical and topographic criteria that the four radiological variables of the model do not capture. Only 3 of these 24 patients were ultimately operated on. The 49 false positives (score ≥ 2 without a retained indication) presented, on the contrary, signs of repercussion (mass effect in 33, impaired consciousness in 34), but had been turned down at the end of the overall assessment. The model is therefore least reliable precisely where the decision escapes the mechanical repercussion visible on CT—first among them malignant infarction—which clearly delimits its scope of application.

4. Discussion

4.1. Main Findings

From a prospective cohort of 250 patients at Bouaké University Hospital (Côte d'Ivoire), we developed and internally validated a parsimonious tool estimating the probability that a patient meets a neurosurgical indication in stroke. Four variables available at admission suffice: hydrocephalus, mass effect and intraventricular extension on CT, and impaired consciousness on examination. They provide moderate discrimination after internal validation reproducing the selection (apparent AUC: 0.77; corrected: 0.73), and overall satisfactory calibration (corrected slope: 0.84; 95% CI: 0.65 - 1.03). The nomogram and its integer-score version offer two modes of bedside use. The decision tree provides a didactic reading of it.

4.2. Pathophysiological Interpretation

The four determinants converge on a single logic: the acute mechanical repercussion of the lesion and obstruction of cerebrospinal fluid circulation. Hydrocephalus, the heaviest marker (100 points), signals an obstruction of the outflow pathways, which frequently justifies emergency drainage. It is a direct and rapidly correctable threat, which explains why it dominates the model. Intraventricular extension is its haemorrhagic corollary. Mass effect reflects intracranial hypertension, which calls for decompression or evacuation. Impaired consciousness is its clinical expression.

The haematoma volume, the Glasgow Coma Scale and the composite scores (ICH, NIHSS) were nevertheless associated with the indication in univariate analysis; they disappear, however, from the final model. This does not reflect a lack of relevance, but redundancy. A haematoma does not mandate a procedure by its size, but by the mass effect and herniation it produces. Likewise, impaired consciousness is better summarised, for the decision, by a binary variable than by the continuous Glasgow Coma Scale. The model thus favours correctable repercussion over overall severity—which is precisely the logic of the neurosurgical proce-

ture, and not that of prognosis (**Figure 11**).

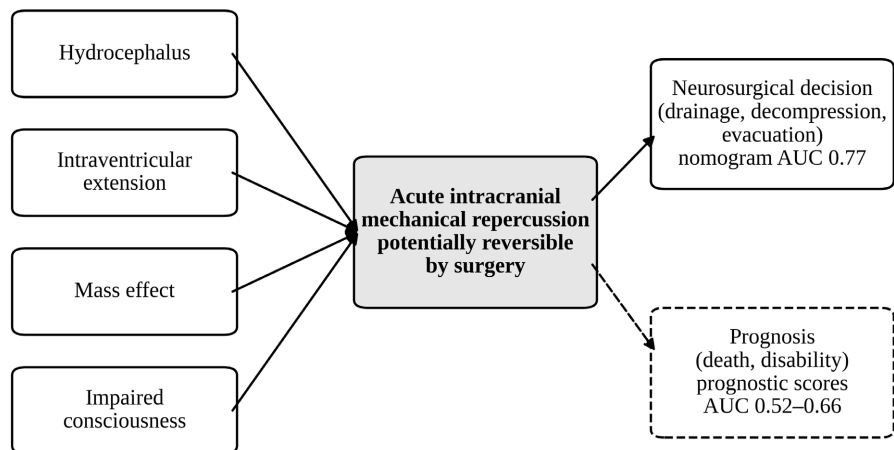


Figure 11. Pathophysiological framework of the neurosurgical decision. The four variables of the model are the manifestations of a single phenomenon—intracranial mechanical repercuSSION potentially reversible by surgery—which distinguishes the operative decision from prognosis.

4.3. Parsimony of the Model

The model rests on four binary variables, accessible at admission. This choice results from statistical selection, but also meets a practical requirement: a tool intended for the emergency room of a resource-limited hospital must be applicable from memory, without calculation or computer support. The trade-off is favourable: rounding into an integer score and choosing its threshold added no measurable optimism to that of the continuous model, optimism remains contained, and two machine-learning methods tested on the same variables did no better. This parsimony guarantees the transparency, interpretability and transposability of the instrument.

4.4. Comparison with the Literature and Guidelines

Our determinants overlap with the criteria that structure the trials and international guidelines: hydrocephalus and intraventricular extension as indications for drainage, mass effect and impaired consciousness as indications for decompression or evacuation [13]–[19] [25]–[29]. These trials chiefly clarified when surgery improves functional outcome. They do not provide a decision-support tool transposable to constrained settings, and they presuppose surgical access that is not guaranteed in our setting [13]–[16] [18] [19]. Nomograms are in common use in intracerebral haemorrhage, but all target prognosis—mortality, functional outcome, sometimes long-term survival to inform the decision indirectly [30] [31]. None takes surgical candidacy as its target: we do not predict what will become of the patient, but what the team will conclude from its assessment.

The comparison with existing scores highlights the specificity of our objective. The classical prognostic scores—ICH score, NIHSS, Glasgow, WFNS, modified Fisher—discriminated the surgical indication only weakly (AUC: 0.52 to 0.66), far

below the nomogram (0.77). This is not a defect of these scores, but the consequence of their purpose: they quantify a risk of death or disability, not candidacy for a procedure. The surgical decision follows a distinct logic, that of correctable mechanical repercussion, which only a dedicated model captures effectively. This finding justifies developing tools specific to the decision, rather than reusing prognostic scores.

4.5. Practical Implications

Several uses emerge. First, support for triage and referral in peripheral hospitals without a neurosurgeon: a high score identifies early the patients with a high probability of indication, to be referred without delay to the neurosurgical facility, and helps to prioritise transfers when transport or admission capacity is limited. Next, a support for harmonising indications between practitioners and for training junior neurosurgeons, useful where expertise is scarce and clinical exposure heterogeneous. Finally, a tool for quantifying surgical need at the scale of a healthcare system, which makes it possible to objectify, as here, the extent of unmet indications. The integer score (0 - 5), which requires no computer support, is particularly suited to use in the emergency room (**Figure 12**).

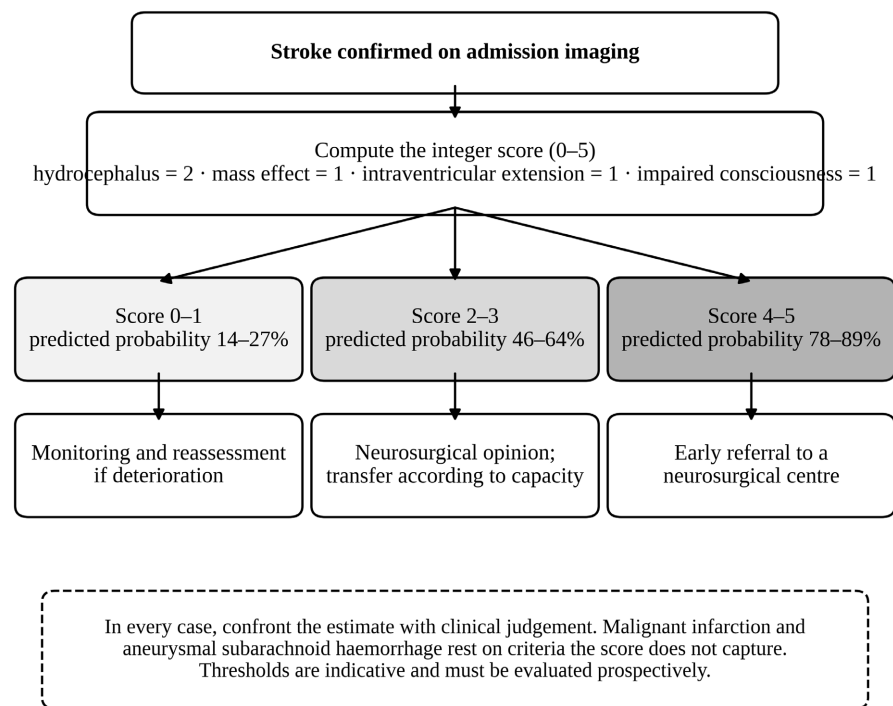


Figure 12. Proposed clinical-use algorithm. The thresholds are indicative, to be evaluated prospectively; they do not replace clinical judgement.

No online calculator is currently available; the model equation (Section 3.3) and the nomogram (**Figure 2**) allow a manual bedside computation.

These uses can be read from a situational grid (**Table 9**); two worked examples follow.

Table 9. Situational grid for use of the score.

Situation	Use of the score	Expected contribution
Peripheral hospital without a neurosurgeon	Identify at admission the patients with a high probability of indication	Early referral and prioritisation of transfers when capacity is limited
Neurosurgery department	Compare the score estimate with clinical judgement	Harmonisation of indications between practitioners, training support
Scale of a department or a territory	Count the indications recorded and those actually met	Reproducible measurement of unmet surgical need, advocacy and planning

Two examples illustrate use of the score. A 58-year-old patient presenting with an intracerebral haemorrhage with hydrocephalus, mass effect and impaired consciousness totals 4 points, *i.e.*, a probability of indication of about 78%: in a peripheral hospital, such a score justifies immediate transfer to a neurosurgical centre. A 62-year-old patient presenting with a limited infarction without hydrocephalus, without mass effect, without intraventricular extension or impaired consciousness totals 0 points, *i.e.*, a probability of about 14%: close monitoring is reasonable, with reassessment in case of deterioration.

4.6. Access Gap to Surgery and Global-Health Implications

The most striking result goes beyond the model's performance: fewer than one indication in two was followed by an intervention. Nearly six patients in ten meeting a surgical indication did not receive it, and all aneurysmal indications remained without therapeutic response. This gap fits the global picture of a massively unmet neurosurgical need in low- and middle-income countries, where several million essential operations are missing each year [32]-[34]. Several non-exclusive factors explain this gap in our setting: the absence of interventional neuroradiology and of dedicated vascular surgery (which explains the total gap for subarachnoid haemorrhage), the limited availability of operating theatres and equipment, the shortage of specialised human resources (neurosurgeons, neuro-anaesthetists), the patients' financial constraints and transfer delays. The respective share of these causes could not be documented patient by patient in the present cohort, which is a limitation; their prospective analysis is a direct perspective of this work.

This gap gives the work a reach that extends beyond individual decision support. By objectifying, patient by patient, the existence of an indication, the tool provides a reproducible instrument for measuring unmet surgical need at the scale of a department or a territory. This is data useful for advocacy, health planning and the sizing of technical facilities. A predictive model cannot fill this gap, but it reveals its extent and allows it to be documented in a standardised way.

4.7. Strengths and Limitations

The strengths of this work lie in the prospective nature of the cohort, the absence

of missing data on the target and on the final predictors—none of the variables with missing data having survived selection—and adherence to the TRIPOD recommendations. Added to this is an internal validation that reproduces the entire development process—imputation, selection and estimation—in each resample: this is the only way to estimate honestly the optimism generated by stepwise selection. An assessment of calibration going beyond the Hosmer-Lemeshow test alone, and a clinical-benefit analysis, complete these strengths.

Several limitations must be underscored. The target reflects the local team's decision. It may incorporate some subjectivity, as well as partial circularity with the radiological variables; it does not constitute an external gold standard, and its inter-observer reproducibility was not measured. Conversely, the fact that a model built on the very criteria that underlie the indication does not exceed 0.77 shows that an appreciable share of the decision escapes the explicit radiological markers. Validation remains purely internal and does not guarantee the transportability of the model, which will have to be established by external multicentre validation before any clinical use. The single-centre nature exposes it to a selection bias: recruited in a referral department, the cohort may over-represent severe forms, and the observed frequency of indication (42.8%) is not transposable to other levels of care. The size limits the granularity of subgroup analyses, notably for subarachnoid haemorrhage and malignant infarction. Some variables were unusable (location, lesion territory), and others were defined only for a subset of the cohort, so that the tree thresholds partly separate haemorrhages from infarctions rather than a gradient of severity. Finally, the study does not demonstrate that use of the model improves decisions or prognosis: it establishes that it predicts an indication, not that it modifies the outcome. The CART tree, presented for illustrative purposes, should not be used in isolation.

A self-assessment of the risk of bias according to the PROBAST tool is proposed in **Table 10** [35].

Table 10. Self-assessment of the risk of bias according to PROBAST.

Domain	Risk of bias	Applicability	Rationale
Participants	Low	Low	Prospective cohort of consecutive patients, explicit inclusion criteria
Predictors	Low	Low	All available at admission and collected before the operative decision; no structured radiological protocol and no double reading
Target	Moderate	Moderate	Target based on the team's expertise, inter-observer reproducibility not measured, partial circularity with the radiological predictors
Analysis	Moderate	—	Number of events per variable of 5.9 relative to the candidates; variables coded at zero outside their domain of definition; CART tree without independent estimation; internal validation, by contrast, reproducing the entire selection

Continued

Overall	Moderate	Low to moderate	Transparent and reproducible model, but developed in a single centre and not externally validated
---------	----------	-----------------	---

4.8. Perspectives

The priority is external multicentre validation, ideally West African, with possible recalibration. In line with the PROBAST and TRIPOD + AI frameworks, it will have to examine explicitly the risk of bias, the transportability and the generalisability of the model [20] [35] [36]. Incorporating the admission delay could enrich the tool. A deployment as an application, coupled with monitoring of unmet indications, would make it possible to combine decision support with steering of the care offer. A prospective evaluation of the score's impact on real decisions would constitute the ultimate step before wide adoption.

5. Conclusion

In this West African neurosurgical cohort, a parsimonious four-variable clinico-radiological model estimates, reproducibly and with overall satisfactory calibration, the probability that a patient is a candidate for a neurosurgical intervention in stroke. To this end, it outperforms the existing prognostic scores. Available as a nomogram and an integer score usable at the bedside, it could structure triage, referral and the prioritisation of transfers in resource-limited settings. It reveals at the same time, a large access gap to surgery—nearly six indications in ten unmet—that no model can fill on its own, but that it makes it possible to objectify. Its external multicentre validation constitutes the indispensable next step.

Funding

This study received no specific funding.

Author Contributions

Role	Authors
Conceptualisation and methodology	All authors
Data collection	Koffi Yves Soress Dongo and Kouadio Jean-Baptiste Keke
Formal analysis	Yao Bernard Fionko
Writing-original draft	Koffi Yves Soress Dongo and Faozo Stephane Landry Teti
Critical revision	All authors
Supervision	Aderehime Haidara

Consent for Publication

The manuscript contains no individual identifying data; consent for publication was not required.

Data Availability

The anonymised data and the analysis code are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] GBD 2019 Stroke Collaborators (2021) Global, Regional, and National Burden of Stroke and Its Risk Factors, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *The Lancet Neurology*, **20**, 795-820. [https://doi.org/10.1016/s1474-4422\(21\)00252-0](https://doi.org/10.1016/s1474-4422(21)00252-0)
- [2] Feigin, V.L., Brainin, M., Norrving, B., Martins, S., Sacco, R.L., Hacke, W., *et al.* (2022) World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. *International Journal of Stroke*, **17**, 18-29. <https://doi.org/10.1177/17474930211065917>
- [3] Johnson, C.O., Nguyen, M., Roth, G.A., Nichols, E., Alam, T., Abate, D., *et al.* (2019) Global, Regional, and National Burden of Stroke, 1990-2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, **18**, 439-458. [https://doi.org/10.1016/s1474-4422\(19\)30034-1](https://doi.org/10.1016/s1474-4422(19)30034-1)
- [4] Kim, J., Thayabaranathan, T., Donnan, G.A., Howard, G., Howard, V.J., Rothwell, P.M., *et al.* (2020) Global Stroke Statistics 2019. *International Journal of Stroke*, **15**, 819-838. <https://doi.org/10.1177/1747493020909545>
- [5] O'Donnell, M.J., Chin, S.L., Rangarajan, S., Xavier, D., Liu, L., Zhang, H., *et al.* (2016) Global and Regional Effects of Potentially Modifiable Risk Factors Associated with Acute Stroke in 32 Countries (INTERSTROKE): A Case-Control Study. *The Lancet*, **388**, 761-775. [https://doi.org/10.1016/s0140-6736\(16\)30506-2](https://doi.org/10.1016/s0140-6736(16)30506-2)
- [6] Owolabi, M.O., Sarfo, F., Akinyemi, R., Gebregziabher, M., Akpa, O., Akpalu, A., *et al.* (2018) Dominant Modifiable Risk Factors for Stroke in Ghana and Nigeria (SIREN): A Case-Control Study. *The Lancet Global Health*, **6**, e436-e446. [https://doi.org/10.1016/s2214-109x\(18\)30002-0](https://doi.org/10.1016/s2214-109x(18)30002-0)
- [7] Adoukonou, T., Kossi, O., Fotso Mefo, P., Agbétou, M., Magne, J., Gbaguidi, G., *et al.* (2021) Stroke Case Fatality in Sub-Saharan Africa: Systematic Review and Meta-Analysis. *International Journal of Stroke*, **16**, 902-916. <https://doi.org/10.1177/1747493021990945>
- [8] Adeloye, D. (2014) An Estimate of the Incidence and Prevalence of Stroke in Africa: A Systematic Review and Meta-Analysis. *PLOS ONE*, **9**, e100724. <https://doi.org/10.1371/journal.pone.0100724>
- [9] Akinyemi, R.O., Ovbiagele, B., Adeniji, O.A., Sarfo, F.S., Abd-Allah, F., Adoukonou, T., *et al.* (2021) Stroke in Africa: Profile, Progress, Prospects and Priorities. *Nature Reviews Neurology*, **17**, 634-656. <https://doi.org/10.1038/s41582-021-00542-4>
- [10] Owolabi, M., Akarolo-Anthony, S., Akinyemi, R., Arnett, D., Gebregziabher, M., Jenkins, C., *et al.* (2015) The Burden of Stroke in Africa: A Glance at the Present and a Glimpse into the Future: Review Article. *Cardiovascular Journal of Africa*, **26**, S27-S38. <https://doi.org/10.5830/cvja-2015-038>
- [11] Akinyemi, R., Sarfo, F., Abd-Allah, F., Ogun, Y., Belo, M., Francis, P., *et al.* (2021) Conceptual Framework for Establishing the African Stroke Organization. *International Journal of Stroke*, **16**, 93-99. <https://doi.org/10.1177/1747493019897871>

- [12] Sarfo, F.S., Akpa, O.M., Ovbiagele, B., Akpalu, A., Wahab, K., Obiako, R., *et al.* (2023) Patient-Level and System-Level Determinants of Stroke Fatality across 16 Large Hospitals in Ghana and Nigeria: A Prospective Cohort Study. *The Lancet Global Health*, **11**, e575-e585. [https://doi.org/10.1016/s2214-109x\(23\)00038-4](https://doi.org/10.1016/s2214-109x(23)00038-4)
- [13] Mendelow, A.D., Gregson, B.A., Fernandes, H.M., Murray, G.D., Teasdale, G.M. and Hope, D.T., *et al.* (2005) Early Surgery versus Initial Conservative Treatment in Patients with Spontaneous Supratentorial Intracerebral Haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): A Randomised Trial. *The Lancet*, **365**, 387-397. [https://doi.org/10.1016/s0140-6736\(05\)70233-6](https://doi.org/10.1016/s0140-6736(05)70233-6)
- [14] Mendelow, A.D., Gregson, B.A., Rowan, E.N., Murray, G.D., Gholkar, A. and Mitchell, P.M. (2013) Early Surgery versus Initial Conservative Treatment in Patients with Spontaneous Supratentorial Lobar Intracerebral Haematomas (STICH II): A Randomised Trial. *The Lancet*, **382**, 397-408. [https://doi.org/10.1016/s0140-6736\(13\)60986-1](https://doi.org/10.1016/s0140-6736(13)60986-1)
- [15] Hanley, D.F., Thompson, R.E., Rosenblum, M., Yenokyan, G., Lane, K., McBee, N., *et al.* (2019) Efficacy and Safety of Minimally Invasive Surgery with Thrombolysis in Intracerebral Haemorrhage Evacuation (MISTIE III): A Randomised, Controlled, Open-Label, Blinded Endpoint Phase 3 Trial. *The Lancet*, **393**, 1021-1032. [https://doi.org/10.1016/s0140-6736\(19\)30195-3](https://doi.org/10.1016/s0140-6736(19)30195-3)
- [16] Pradilla, G., Ratcliff, J.J., Hall, A.J., Saville, B.R., Allen, J.W., Paulon, G., *et al.* (2024) Trial of Early Minimally Invasive Removal of Intracerebral Hemorrhage. *New England Journal of Medicine*, **390**, 1277-1289. <https://doi.org/10.1056/nejmoa2308440>
- [17] Beck, J., Fung, C., Strbian, D., Bütikofer, L., Z'Graggen, W.J., Lang, M.F., *et al.* (2024) Decompressive Craniectomy Plus Best Medical Treatment versus Best Medical Treatment Alone for Spontaneous Severe Deep Supratentorial Intracerebral Haemorrhage: A Randomised Controlled Clinical Trial. *The Lancet*, **403**, 2395-2404. [https://doi.org/10.1016/s0140-6736\(24\)00702-5](https://doi.org/10.1016/s0140-6736(24)00702-5)
- [18] Vahedi, K., Hofmeijer, J., Juettler, E., Vicaut, E., George, B., Algra, A., *et al.* (2007) Early Decompressive Surgery in Malignant Infarction of the Middle Cerebral Artery: A Pooled Analysis of Three Randomised Controlled Trials. *The Lancet Neurology*, **6**, 215-222. [https://doi.org/10.1016/s1474-4422\(07\)70036-4](https://doi.org/10.1016/s1474-4422(07)70036-4)
- [19] Molyneux, A.J., Kerr, R.S., Yu, L., Clarke, M., Sneade, M., Yarnold, J.A., *et al.* (2005) International Subarachnoid Aneurysm Trial (ISAT) of Neurosurgical Clipping versus Endovascular Coiling in 2143 Patients with Ruptured Intracranial Aneurysms: A Randomised Comparison of Effects on Survival, Dependency, Seizures, Rebleeding, Subgroups, and Aneurysm Occlusion. *The Lancet*, **366**, 809-817. [https://doi.org/10.1016/s0140-6736\(05\)67214-5](https://doi.org/10.1016/s0140-6736(05)67214-5)
- [20] Collins, G.S., Reitsma, J.B., Altman, D.G. and Moons, K.G.M. (2015) Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Annals of Internal Medicine*, **162**, 55-63. <https://doi.org/10.7326/m14-0697>
- [21] Iasonos, A., Schrag, D., Raj, G.V. and Panageas, K.S. (2008) How to Build and Interpret a Nomogram for Cancer Prognosis. *Journal of Clinical Oncology*, **26**, 1364-1370. <https://doi.org/10.1200/jco.2007.12.9791>
- [22] Steyerberg, E.W., Harrell, F.E., Borsboom, G.J.J.M., Eijkemans, M.J.C., Vergouwe, Y. and Habbema, J.D.F. (2001) Internal Validation of Predictive Models: Efficiency of Some Procedures for Logistic Regression Analysis. *Journal of Clinical Epidemiology*, **54**, 774-781. [https://doi.org/10.1016/s0895-4356\(01\)00341-9](https://doi.org/10.1016/s0895-4356(01)00341-9)
- [23] Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984) Classification and Regression Trees. Wadsworth International Group, x+358 p.

- [24] von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C. and Vandenbroucke, J.P. (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *The Lancet*, **370**, 1453-1457. [https://doi.org/10.1016/s0140-6736\(07\)61602-x](https://doi.org/10.1016/s0140-6736(07)61602-x)
- [25] Hemphill, J.C., Bonovich, D.C., Besmertis, L., Manley, G.T. and Johnston, S.C. (2001) The ICH Score: A Simple, Reliable Grading Scale for Intracerebral Hemorrhage. *Stroke*, **32**, 891-897. <https://doi.org/10.1161/01.str.32.4.891>
- [26] Greenberg, S.M., Ziai, W.C., Cordonnier, C., Dowlatshahi, D., Francis, B., Goldstein, J.N., *et al.* (2022) 2022 Guideline for the Management of Patients with Spontaneous Intracerebral Hemorrhage: A Guideline from the American Heart Association/American Stroke Association. *Stroke*, **53**, e282-e361. <https://doi.org/10.1161/str.0000000000000407>
- [27] Steiner, T., Salman, R.A., Beer, R., Christensen, H., Cordonnier, C., Csiba, L., *et al.* (2014) European Stroke Organisation (ESO) Guidelines for the Management of Spontaneous Intracerebral Hemorrhage. *International Journal of Stroke*, **9**, 840-855. <https://doi.org/10.1111/ijss.12309>
- [28] Steiner, T., Purrucker, J.C., Aguiar de Sousa, D., Apostolaki-Hansson, T., Beck, J., Christensen, H., *et al.* (2025) European Stroke Organisation (ESO) and European Association of Neurosurgical Societies (EANS) Guideline on Stroke Due to Spontaneous Intracerebral Haemorrhage. *European Stroke Journal*, **10**, 1007-1086. <https://doi.org/10.1177/23969873251340815>
- [29] Powers, W.J., Rabinstein, A.A., Ackerson, T., Adeoye, O.M., Bambakidis, N.C., Becker, K., *et al.* (2019) Guidelines for the Early Management of Patients with Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke: A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association. *Stroke*, **50**, e344-e418. <https://doi.org/10.1161/str.0000000000000211>
- [30] Gregório, T., Pipa, S., Cavaleiro, P., Atanásio, G., Albuquerque, I., Chaves, P.C., *et al.* (2018) Prognostic Models for Intracerebral Hemorrhage: Systematic Review and Meta-Analysis. *BMC Medical Research Methodology*, **18**, Article No. 145. <https://doi.org/10.1186/s12874-018-0613-8>
- [31] Lin, F., He, Q., Zhuo, L., Zhao, M., Ye, G., Gao, Z., *et al.* (2023) A Nomogram Predictive Model for Long-Term Survival in Spontaneous Intracerebral Hemorrhage Patients without Cerebral Herniation at Admission. *Scientific Reports*, **13**, Article No. 3126. <https://doi.org/10.1038/s41598-022-26176-0>
- [32] Dewan, M.C., Rattani, A., Fieggen, G., Arraez, M.A., Servadei, F., Boop, F.A., *et al.* (2019) Global Neurosurgery: The Current Capacity and Deficit in the Provision of Essential Neurosurgical Care. Executive Summary of the Global Neurosurgery Initiative at the Program in Global Surgery and Social Change. *Journal of Neurosurgery*, **130**, 1055-1064. <https://doi.org/10.3171/2017.11.jns171500>
- [33] Punchak, M., Mukhopadhyay, S., Sachdev, S., Hung, Y., Peeters, S., Rattani, A., *et al.* (2018) Neurosurgical Care: Availability and Access in Low-Income and Middle-Income Countries. *World Neurosurgery*, **112**, e240-e254. <https://doi.org/10.1016/j.wneu.2018.01.029>
- [34] Gupta, S., Gal, Z.T., Athni, T.S., Calderon, C., Callison, W.É., Dada, O.E., *et al.* (2024) Mapping the Global Neurosurgery Workforce. Part 1: Consultant Neurosurgeon Density. *Journal of Neurosurgery*, **141**, 1-9. <https://doi.org/10.3171/2023.9.jns231615>
- [35] Wolff, R.F., Moons, K.G.M., Riley, R.D., Whiting, P.F., Westwood, M., Collins, G.S., *et al.* (2019) PROBAST: A Tool to Assess the Risk of Bias and Applicability of Predic-

tion Model Studies. *Annals of Internal Medicine*, **170**, 51-58.

<https://doi.org/10.7326/m18-1376>

- [36] Collins, G.S., Moons, K.G.M., Dhiman, P., Riley, R.D., Beam, A.L., Van Calster, B., *et al.* (2024) TRIPOD + AI Statement: Updated Guidance for Reporting Clinical Prediction Models That Use Regression or Machine Learning Methods. *BMJ*, **385**, e078378. <https://doi.org/10.1136/bmj-2023-078378>

Abbreviations

aOR = adjusted odds ratio;

AUC = area under the ROC curve;

CART = classification and regression tree;

CI = confidence interval;

DCA = decision curve analysis;

EVD = external ventricular drain;

ICH = intracerebral hemorrhage (score);

IQR = interquartile range;

LR+ = positive likelihood ratio;

NIHSS = National Institutes of Health Stroke Scale;

NPV = negative predictive value;

PPV = positive predictive value;

PROBAST = Prediction model Risk of Bias ASsessment Tool;

SAH = subarachnoid haemorrhage;

STROBE = Strengthening the Reporting of Observational Studies in Epidemiology;

TRIPOD = Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis;

VIF = variance inflation factor;

WFNS = World Federation of Neurosurgical Societies (grade).