Clinical risk scores for predicting stroke-associated pneumonia: a systematic review

Amit K Kishore[1], Andy Vail[2] , Benjamin D Bray[3], Angel Chamorro[4], Mario Di Napoli[5], Lalit Kalra[6], Peter Langhorne[7], Joan Montaner[8,9], Christine Roffe[10], Anthony G Rudd[11], Pippa J Tyrrell[1], Diederik van de Beek[12], Mark Woodhead[13], Andreas Meisel[14], Craig J Smith[1]

[1]Stroke and Vascular Research Centre, University of Manchester, Institute of Cardiovascular Sciences, Greater Manchester Comprehensive Stroke Centre Salford Royal Foundation Trust, UK

[2]Centre for Biostatistics, University of Manchester, Salford Royal Foundation Trust, UK

[3]King's College, London, UK

[4]Comprehensive Stroke Centre, Department of Neuroscience, Hospital Clinic, University of Barcelona, Barcelona, Spain;

[5]Neurological Service, San Camillo de' Lellis General Hospital, Rieti, Italy

[6]Clinical Neurosciences, King's College Hospital NHS Foundation Trust London, UK

[7]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow Royal Infirmary, Glasgow, UK

[8]Laboratorio de Investigación Neurovascular, Unidad Neurovascular, Servicio de Neurología Hospital Vall d' Hebron, Barcelona

[9]IBIS Stroke Programme, Hospital Virgen del Rocio, Sevilla, Spain

[10]Keele University Institute for Science and Technology in Medicine, Guy Hilton Research Centre, Stoke-on-Trent, UK

[11]Department of Health and Social Care, Kings College, London, UK

[12]Department of Neurology, Centre for Infection and Immunity Amsterdam (CINIMA), Academic Medical Centre, University of Amsterdam, Amsterdam, Netherlands

[13]Faculty of Medical and Human Sciences, University of Manchester & Department of Respiratory Medicine, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK

[14]NeuroCure Clinical Research Centre, Centre for Stroke Research Berlin, Department of Neurology Charité Universitaetsmedizin Berlin, Germany

Corresponding author:
Dr Amit K Kishore, Greater Manchester Neurosciences Centre, Salford Royal NHS Foundation Trust, Stott lane, Salford, M6 8HD, UK.
E-mail: Amit.Kishore@manchester.ac.uk

Cover title: Prediction scores for stroke-associated pneumonia

Keywords: pneumonia, ischaemic stroke, intracerebral haemorrhage, risk score

Purpose

Several risk stratification scores for predicting stroke-associated pneumonia (SAP) have been derived. We aimed to evaluate the performance and clinical usefulness of such scores for predicting SAP.

Method

A systematic literature review was undertaken in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, with application of the Quality Assessment of Diagnostic Accuracy (QUADAS)-2 tool. Published studies of hospitalised adults with ischaemic stroke, intracerebral haemorrhage, or both, which derived and validated an integer-based clinical risk score, or externally validated an existing score to predict occurrence of SAP, were considered and independently screened for inclusion by two reviewers.

Findings

We identified 9 scores, from 8 derivation cohorts. Age was a component of all scores, and the NIHSS score in all except one. 6 scores were internally validated and 5 scores were externally validated. The A2DS2 score (Age, Atrial fibrillation, Dysphagia, Severity [NIHSS], Sex) was the most externally validated in 8 independent cohorts. Performance measures were reported for 8 scores. Discrimination tended to be more variable in the external validation cohorts (C statistic 0.67-0.83) than the derivation cohorts (C statistic 0.74-0.85).

Discussion

Overall, discrimination and calibration were similar between the different scores. No study evaluated influence on clinical decision-making or prognosis.

Conclusion

The clinical prediction scores varied in their simplicity of use and were comparable in performance. Utility of such scores for preventive intervention trials and in clinical practice remains uncertain and requires further study.

Introduction

Stroke-associated pneumonia (SAP) is a common and serious complication after acute stroke, associated with increased length of hospital stay, mortality and worse outcomes in survivors.[1-6] A recent systematic review reported that SAP occurs in 14.3%, although the frequency varies widely depending on definition of SAP and patient characteristics.[7] Several features of SAP such as varied clinical manifestation,[8] uncertain role of blood biomarkers[9] and absence of definitive diagnostic criteria make it challenging to diagnose in clinical practice. As a first step, the recently convened Pneumonia In Stroke ConsEnsuS (PISCES) group proposed operational diagnostic criteria for SAP based on Center of Disease Control criteria (CDC).[9]

Numerous baseline clinical factors such as age, dysphagia, severity of stroke, low conscious level, type and location of stroke may pre-dispose individuals to SAP.[10-13] Predictive risk models derived using these routinely available variables may help in identifying patients at an increased risk of pneumonia for targeted preventive measures and may also provide opportunities for novel interventions for monitoring or therapy. However, clinical prediction scores have several potential weaknesses such as differences in derivation, inconsistent external validation and complexity thus making choice of score and application to clinical practice challenging.[14,15] We therefore undertook a systematic review to identify scores used in predicting risk of SAP, with the aim of evaluating performance, usability and utility for clinical practice and research.

Methods

A systematic literature review was undertaken in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[16]

Data sources and searches

Searches were undertaken in MEDLINE (1946-15th September 2015) and EMBASE (1947-15th September 2015) using pre-defined search criteria and terms (**Online only Table I**). Hand searching of reference lists for additional eligible articles was also carried out, and the PISCES group were invited to provide any other potentially eligible articles.

Study selection

Published studies (English and Non-English) of hospitalised adults with ischaemic stroke, intracerebral haemorrhage (ICH), or both, which derived and validated an integer-based clinical risk score, or externally validated an existing score to predict occurrence of pneumonia after stroke, were independently screened for eligibility by two reviewers (AKK and CJS), using the study title and abstract (**Online Only Table II**). Lead or corresponding authors of studies under consideration were contacted by e-mail to resolve any issues relating to assessment of eligibility or data extraction. Discrepancies relating to eligibility or data extraction were resolved by discussion between the same two study investigators.

Data Extraction

Data were independently extracted by two reviewers (AKK and CJS) and included study design, clinical environment, country, stroke subtype (ischaemic or ICH), mean age, mean National Institutes of Health Stroke Scale (NIHSS) score, components of score and

weighting, measures of discrimination and calibration, co-morbidities, criteria used in diagnosis of pneumonia and proportion of patients diagnosed with pneumonia.

Assessment of quality: risk of bias and applicability

Quality was assessed in terms of risk of bias and concerns regarding applicability, using the Quality Assessment of Diagnostic Accuracy (QUADAS)-2 tool.[17] In brief, judgement of applicability and risk of bias are made across 4 domains using relevant signalling questions; patient selection, index risk score, reference standard (diagnosis of SAP) and flow and timing. The QUADAS-2 tool was applied for each score within the identified validation cohorts by two reviewers (AKK and AV) independently.

Risk score performance

For the discriminative ability of scores, we extracted information on the area under the receiver operating characteristic curve (AUROC) or C-statistic, their 95% confidence intervals, and the p-value for comparison between models if they were available for both the derivation and validation cohorts. C-statistic values range from 0.5 (no discrimination, no better than chance) to 1.0 (perfect discrimination). A C-statistic of 0.7-0.8 indicates modest discriminative ability, while a C-statistic greater than 0.8 indicates good discriminative ability. To describe score calibration, we similarly extracted data on the difference between the observed and predicted rates of pneumonia if available, as well as the 'goodness of fit' statistic and p-value of the corresponding test statistic. Calibration was considered better when the observed to predicted ratio was closer to 1.

Clinical usefulness

We noted the complexity of application and use at the bedside, whether prediction scores incorporated categories of risk-stratification (usability), and whether scores had been used to evaluate clinical management or clinician behaviors (utility). We also evaluated the generalisability of each prediction model by determining whether it had been externally validated in an independent patient population, either in the original or subsequent publication.

Findings

Search results

The electronic search yielded 2493 publications. After screening, exclusion of duplicates and applying eligibility criteria, 46 full texts and abstracts were reviewed **(Figure 1).** No additional articles were identified through hand-searching of major stroke journals or by the PISCES group. 12 fully published studies were finally considered eligible for inclusion.[11, 18-28]

Clinical risk scores for predicting SAP

14 separate cohorts were identified which had either derived or validated risk scores for predicting SAP **(Online only Table III)**. 9 clinical risk-scores (**Table 1**) were identified from 8 derivation cohorts. The risk scores identified were: The Pneumonia Score,[18] Veteran's Health Administration cohort score,[11] A2DS2 (Age, Atrial fibrillation, Dysphagia, Severity [NIHSS], Sex) score,[19] PANTHERIS (Preventive Antibacterial Therapy in Acute Ischaemic Stroke) score,[20] AIS-APS (Acute Ischaemic Stroke-Associated Pneumonia Score),[21] ICH-APS (Intracerebral Haemorrhage-Associated Pneumonia Score) A,[22] ICH-APS score B,[22] Pneumonia (PNA) prediction score[25] and ISAN (Prestroke Independence, Sex, Age, NIHSS) score.[27] The number of components of each score varied between 3 and 11. Age was a component of all scores, and the NIHSS score in all except the PANTHERIS score.[20] Dysphagia (55%) and pre-morbid modified Rankin Scale (mRS, 36%) were other commonly used variables. 2 scores also included routine laboratory evaluation[20,21] and 2 scores incorporated neuroimaging features.[22] The majority of the scores were derived only in ischaemic stroke patients (56%). The ISAN and the PNA prediction score included ICH patients in the derivation cohorts, [25, 27] while the ICH-APS scores[22] were developed exclusively in ICH.

Study and participant characteristics

Characteristics of the derivation and validation cohorts are summarised in **Online only Table III**. Median age was 71y (range 61y-76y) and median NIHSS was 5 (range 4-13). All studies adequately described selection of study sample. Except one study which was prospective, [26] all were retrospective evaluations of existing prospective cohorts. Of the 14 separate derivation or validation cohorts, 8 (61%) were multicentre or national stroke registries and 6 (39%) were single-centre hospital-based stroke registries. The majority of the 14 studies (80%) evaluated only acute ischaemic stroke. Definition and ascertainment of risk factors for model derivation was varied and often limited by availability of data, particularly in existing national registries. For example, dysphagia assessment was not described among several studies, [21, 22, 25] and pre- stroke disability was described in different ways. [21,22,27] Some studies did not record pre-existing disability.[18,25] Diagnostic approach to pneumonia varied between cohorts; clinician reported diagnosis of pneumonia (36%) and the CDC criteria for pneumonia (36%) were the most commonly used approaches. The other methods include adhoc objective criteria (14%) and Chinese Consensus criteria (14%).

Quality assessment

Overall, risk of bias and concerns regarding applicability were judged as generally low (Online only **Table IV)**. In some validation cohorts, risk of bias was judged as high based on patient selection (exclusions based on incomplete baseline data[27], or selected higher-risk cohort[20]), reference standard (non-standardised criteria for diagnosis of SAP[11,19,27,28]) and flow and timing (verification bias, related to differences in applying the same reference standard by the study group[11,19,27,28]).

Performance and validation of the risk-scores

Performance and validation of the clinical risk scores is summarized in **Table 2** and **3**. C-statistics ranged from 0.74-0.85 in the derivation cohorts, indicating a degree of discriminative performance that varied from modest to good. Only one score, the Pneumonia score, [18] did not report performance. Calibration was reported in 6 models using varying goodness of fit models or net reclassification index.

6 of the risk-scores were validated internally through split samples (**Table 2**). All reported the C-statistic for the internal validation cohort, which ranged from 0.73 to 0.88, with 5 models reporting calibration metric. 5 of the 9 scores were validated externally (**Table 3**), with C-statistic ranging from 0.68 to 0.83. The A2DS2 score[19] has been evaluated most extensively, in the largest derivation sample (n=15,335), and in 8 separate external validation cohorts. The A2DS2 score performed consistently across these cohorts (C statistic 0.73 to 0.84), with good calibration.

Clinical usefulness

The risk-scores varied in their complexity and ease of use (**Table 1**), although most scores incorporated clinical variables readily available at baseline. 2 scores require admission laboratory variables[20,21] and one of the scores developed exclusively for ICH requires quantitative measurement of hematoma volume.[22] Several of the scores were stratified into integer-based risk categories[11,19,21,22,27] (e.g. low, moderate, high risk), facilitating usability by clinicians. The role of implementing the models as prediction rules in terms of risk stratification, decision-making or improved patient outcomes was not evaluated for any of the scores.

Discussion

An ideal risk score for predicting SAP would incorporate variables readily available at stroke presentation, be quick to apply, provide meaningful risk categories with performance acceptable to the particular application (and to clinicians), and have impact on clinical decision-making and clinical outcomes. In this systematic review, we identified nine clinical risk scores for predicting SAP, and assessed their performance metrics, clinical usability and utility. We sought to identify whether any of the scores could be applied for use in clinical care or research.

The scores varied considerably in their complexity, component variables, derivation cohort characteristics, approach to defining SAP, ease of application, consistency of external validation, and performance evaluation. Substantial heterogeneity between the studies was therefore anticipated and precluded meta-analyses. As previously acknowledged, [27] the prevalence of SAP varied between the cohorts, most likely related to underlying differences in patient characteristics and definitions used, [1] potentially contributing to outcome reporting bias. Several of the scores were derived from relatively small single centre cohorts[18, 20, and 25] limiting their generalisability. As all of the scores were derived using retrospective analyses of registry-based studies, model-building was limited by the baseline characteristics recorded in the different cohorts. Therefore, potentially important baseline characteristics (e.g. smoking, medication, chronic lung disease), medications (e.g. statin therapy or beta-blockade), laboratory variables (e.g. leukocyte count or C-reactive protein [CRP]) or interventions (e.g. mechanical intervention, type of swallow screen), which may have influenced SAP risk, were not available in the majority of the derivation cohorts.

For the studies reporting performance metrics, the discriminative ability and calibration of the scores ranged from moderate to good. However, several of the scores have not yet undergone external validation to our knowledge.[20,22,25] Some of the scores performed similarly in the external validation and derivation cohorts,[19,21,27] despite differences in patient characteristics, supporting generalisability. Importantly, the majority of the validation studies were unable to compare the performance of more than one score concurrently due to limitations imposed by data routinely collected in the registry-based cohorts. One study compared four scores concurrently, [21] and found no material difference in the performance metrics of the four scores (Pneumonia score, VHA score, AIS-APS, A2DS2) tested. Most scores were derived only in ischaemic stroke cohorts, although 2 scores with comparable performance were available for ICH. The ISAN and A2DS2 were evaluated in both ischaemic stroke and ICH, and performance metrics tended to be superior in ischaemic stroke rather than ICH, most likely due to ceiling effects.[27] The only scores derived exclusively for ICH (ICH-APS A and B) are less practical to apply, requiring baseline imaging parameters, and have not been externally validated to date.[22] Considering the high-rate of early neurological deterioration and conflicting risk of death after ICH, the ISAN, A2DS2 and ICH-APS scores each performed better, and comparably, in sensitivity analyses stratifying for survival beyond 48-72 h after ICH.[22,27]

The role of clinical risk-scores for predicting SAP in clinical care or research remains uncertain. None of the studies investigated utility in terms of clinician behaviours (for example, the time taken to administer the risk scores) or impact analysis on clinical outcomes. The current levels of sensitivity and specificity for given cut-offs on the scores[19, 21, 22, 24] may be unacceptable to clinicians, although this may depend on the particular application of the score. For example, for a cut-off of ≥4 on the A2DS2 score, sensitivity is

91% but specificity is 57%.[19] This means that only 9% of actual SAP cases are not identified as high-risk (false negative rate), yet 43% of the patients who do not get SAP are incorrectly identified as being at high-risk (false positive rate). For a safe, inexpensive and well-tolerated intervention to prevent SAP (e.g. enhanced monitoring or oral hygiene protocol) this extent of exposure to unnecessary interventions may be acceptable. However, for more expensive and complex preventive interventions with adverse effects, which are challenging to administer, then such low specificity may make clinical trials impractical and more difficult to justify.

Further large, multi-centre prospective studies of consecutive patients, with adjudicated diagnosis of SAP using standardised and validated criteria are required to evaluate comparative performance and utility of the available scores. Refining the existing scores, including the addition of laboratory biomarkers such as CRP to improve performance, [26] warrants further consideration. Finally, evaluating clinical utility of the scores is an essential step to determine effects on clinician behaviours, impact on clinical decision-making, clinical outcomes and feasibility of implementation.

Conclusion

We identified several clinical risk scores for predicting SAP which varied in their simplicity and consistency of validation. When recorded, performance metrics were comparable between scores, and no single score consistently performed better than others. However, interpretation was limited by heterogeneity and some risk of bias. The utility of risk scores for predicting SAP remains uncertain and requires further study in prospective cohorts with standardised criteria for definition of SAP.

References

1. Katzan IL, Cebul RD, Husak SH, et al. The effect of pneumonia on mortality among patients hospitalized for acute stroke. *Neurology* 2003; 60:620–625.

2. Ovbiagele B, Hills NK, Saver JL, et al; for the California Acute Stroke Prototype Registry Investigators. Frequency and determinants of pneumonia and urinary tract infection during stroke hospitalization. *J Stroke Cerebrovasc Dis* 2006; 15: 209–213.

3. Finlayson O, Kapral M, Hall R, et al; on behalf of the Investigators of the Registry of the Canadian Stroke Network. Risk factors, inpatient care, and outcomes of pneumonia after ischemic stroke. *Neurology* 2011; 77: 1338–1345.

4. Katzan IL, Dawson NV, Thomas CL, et al. The cost of pneumonia after acute stroke. *Neurology* 2007; 68: 1938–1943.

5. Koennecke HC, Belz W, Berfelde D, et al; for the Berlin Stroke Register Investigators. Factors influencing in-hospital mortality and morbidity inpatients treated on a stroke unit. *Neurology* 2011; 77: 965–972.

6. Kammersgaard LP, Jorgensen HS, Reith J, et al. Early infection and prognosis after acute stroke: the Copenhagen stroke study. *J Stroke Cerebrovasc Dis* 2001; 10:217–221.

7. Kishore AK, Vail A, Chamorro A, et al. How is pneumonia diagnosed in clinical stroke research? A systematic review and meta-analysis. *Stroke* 2015; 46: 1202-1209.

8. Harms H, Hoffmann S, Malzahn U, et al. Decision-making in the diagnosis and treatment of stroke-associated pneumonia. *J Neurol Neurosurg Psychiatry* 2012; 83: 1225-1230.

9. Smith CJ, Kishore AK, Vail A, et al. Diagnosis of stroke-associated pneumonia: recommendations from the Pneumonia In Stroke ConsEnsuS Group. *Stroke* 2015; 46: 2335-2340.

10. Finlayson O, Kapral M, Hall R, et al. Risk factors, inpatient care, and outcomes of pneumonia after ischemic stroke. *Neurology* 2011;77: 1338-1345.

11. Chumbler NR, Williams LS, Wells CK, et al. Derivation and validation of a clinical system for predicting pneumonia in acute stroke. *Neuroepidemiology* 2010; 34:193–199.

12. Sellars C, Bowie L, Bagg J, et al. Risk factors for chest infection in acute stroke: a prospective cohort study. *Stroke* 2007; 38: 2284–2291.

13. Kemmling A, Lev MH, Payabvash S, et al. Hospital acquired pneumonia is linked to right hemispheric peri-insular stroke. *PLoS One* 2013; 8: e71141.

14. Yealy DM, Auble TE. Choosing between clinical prediction rules. *N Engl J Med* 2003; 349:2553-2555.

15. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation* 2010; 121:1768-77.

16. Moher D, Liberati A, Tetzlaff J, et al. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 2009;6:e1000097.

17. Whiting PF, Rutjes AWS, Westwood ME, et al; and the QUADAS-2 Group. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011; 155: 529-536.

18. Kwon H, Jeong S, Lee S, et al. The pneumonia score: A simple grading scale for prediction of pneumonia after acute stroke. *Am J Infect Control* 2006; 34:64-68.

19. Hoffmann S, Malzahn U, Harms H, et al; Berlin Stroke Register and the Stroke Register of Northwest Germany. Development of a clinical score (A2DS2) to predict pneumonia in acute ischaemic stroke. *Stroke* 2012;43: 2617–2623.

20. Harms H, Grittner U, Droge H, et al. Predicting post-stroke pneumonia: the PANTHERIS score. *Acta Neurol Scand* 2013; 128:178-184.

21. Ji R, Shen H, Pan Y, et al. China National Stroke Registry Investigators. Novel risk score to predict pneumonia after acute ischaemic stroke. Stroke 2013; 44:1303-1309.

22. Ji R, Shen H, Pan Y, et al; China National Stroke Registry Investigators. Risk Score to Predict Hospital-Acquired Pneumonia After Spontaneous Intracerebral Haemorrhage. *Stroke* 2014;45: 2620-2628.

23. Li L, Zhang L, Xu W, et al. Risk assessment of ischaemic stroke associated pneumonia. *World J Emerg Med* 2014; 5:209–213.

24. Li Y, Song B, Fang H, et al. External Validation of the A2DS2 Score to Predict Stroke-Associated Pneumonia in a Chinese Population: A Prospective Cohort Study. *PLoS ONE* 2014; 9(10): e109665.

25. Friedant AJ, Gouse BM, Boehme AK, et al. A simple prediction score for developing a hospital-acquired infection after acute ischaemic stroke. *J Stroke Cerebrovasc Dis* 2015;24: 680-686.

26. Lu Y, Liu X, Chen Y, et al. Serum iron and A(2)DS(2) score in stroke-associated pneumonia. *Int J Clin Exp Med* 2015; 8: 6163-6170.

27. Smith CJ, Bray BD, Hoffman A, et al. Can a novel clinical risk score improve pneumonia prediction in acute stroke care? A UK multicenter cohort study. *J Am Heart Assoc* 2015; 4.

28. Papavasileiou VV, Milionis H, Smith CJ, et al. External Validation of the Prestroke Independence, Sex, Age, National Institutes of Health Stroke Scale (ISAN) Score for Predicting Stroke-Associated Pneumonia in the Athens Stroke Registry. *J Stroke and Cerebrovasc Dis* 2015; 24: 2619-2624.

Table 1: Components of clinical risk scores for predicting stroke-associated pneumonia

	ISAN	A2DS2	AIS-APS	PANTHERIS	VHA cohort	Pneumonia score	PNA prediction score	ICH-APS-A	ICH-APS-B
Age	✓	✓	✓	✓	✓	✓	✓	✓	✓
NIHSS	✓	✓	✓		✓	✓	✓	✓	✓
GCS			✓	✓				✓	
Sex	✓	✓				✓			
Dysphagia		✓			✓	✓		✓	✓
Mechanical ventilation						✓			
Dysphasia			✓						
OCSP subtype			✓						
"Found-down"					✓				
Increase in systolic BP > 200mmHg				✓					
Comorbidities									
Pre-stroke dependence (mRS)	✓		✓					✓	✓
Atrial fibrillation		✓	✓						
Congestive cardiac failure			✓						
COPD			✓					✓	✓
Current smoking			✓					✓	✓
Excess alcohol consumption								✓	✓
Previous pneumonia					✓				
Diabetes							✓		
Laboratory									
Blood glucose (mmol/l)			✓						
WBC count /µl				✓					
Radiology									
Infratentorial location								✓	✓
Extension into ventricles								✓	
Haematoma volume									✓

ISAN indicates Independence Prestroke, Sex, Age, National Institutes of Health Stroke Scale; A2DS2, Age, Atrial Fibrillation, Dysphagia Sex, Severity; AIS-APS, Acute Ischaemic Stroke-Associated Pneumonia Score; PANTHERIS, Preventive Antibacterial Therapy in Acute Ischaemic Stroke; VHA, Veterans Health Administration; PNA, Pneumonia Prediction; ICH-APS, Intracerebral haemorrhage-Associated Pneumonia Score; NIHSS: National Institutes of Health Stroke Scale; GCS: Glasgow coma scale; OCSP: Oxfordshire community stroke project; mRS: modified Rankin scale; COPD: chronic obstructive pulmonary disease; WBC: White blood cell.

Table 2: Performance of the clinical risk scores for predicting SAP in the derivation and internal validation cohorts

Score	Derivation Cohort	Size (n)	Frequency of SAP (%)	SAP diagnostic criteria	Discrimination metric; C statistic (95% CI)	Calibration metric
ISAN score	SSNAP	11551	6.7	Clinician reported	0.79 (0.77 to 0.81)	'well calibrated across all 22 levels in validation cohort; ceiling effect for score > 15 among ICH patients
A2DS2	BSR	15335	7.2	Clinician reported	0.84 (0.83-0.85)	Cox and Snell *R*2 =0.106, Nagelkerke R^2=0.259, McFadden R^2 = 0.213
AIS-APS	CNSR	8820	11.4	CDC	0.80 (0.78–0.81)	NR
PANTHERIS	Berlin NICU	223	30	CDC	0.85 (0.80–0.91)	Nagelkerke's R^2 0.46
VHA score	VHA	925	10.4	Clinician reported	0.76 (NR)	2.1% misclassification
Pneumonia score	Seoul	286	10.5	Adhoc objective	NR	NR
PNA score	New-Orleans	568	11.4	Adhoc objective	0.79 (NR)	NR
ICH-APS(A)	CNSR	2998	17	CDC	0.75 (0.72–0.77)	Hosmer–Lemeshow goodness of fit test, *P*=0.20
ICH-APS (B)	CNSR	2998	17	CDC	0.74 (0.71–0.76)	Hosmer–Lemeshow goodness of fit test, *P*=0.10
Score	Internal Validation Cohort	Size (n)	Frequency of SAP (%)	SAP diagnostic criteria	Discrimination metric; C statistic (95% CI)	Calibration metric
ISAN score	SSNAP	11648	6.7	Clinician reported	0.78 (0.76-0.80)	NR
AIS-APS	CNSR	5882	11.3	CDC	0.79 (0.77–0.80)	The Hosmer–Lemeshow test was not significant *P*=0.22
PANTHERIS	Berlin NICU	112	33.9	CDC	0.88 (0.81–0.95)	Nagelkerke's R^2 = 0.48
VHA score	VHA	438	10.5	Clinician reported	0.78	0.9% reclassification
ICH-APS(A)	CNSR	2000	16.7	CDC	0.76 (0.71–0.79)	Hosmer–Lemeshow goodness of fit test, *P*=0.66
ICH-APS (B)	CNSR	2000	16.7	CDC	0.73 (0.70–0.76	Hosmer–Lemeshow goodness of fit test, *P*=0.17

SAP indicates Stroke-Associated Pneumonia; ISAN, Independence Prestroke, Sex, Age, National Institutes of Health Stroke Scale; A2DS2, Age, Atrial, Fibrillation, Dysphagia, Sex, Severity; AIS-APS, Acute Ischemic Stroke-Associated Pneumonia Score; PANTHERIS, Preventive Antibacterial Therapy in Acute Ischaemic Stroke; VHA, Veterans Health Administration; PNA, Pneumonia Prediction; ICH-APS, Intracerebral Haemorrhage-Associated Pneumonia Score; SSNAP, Sentinel Stroke National Audit Programme; BSR, Berlin Stroke Register; CNSR, Chinese National Stroke Registry; Berlin NICU, Berlin Neurological Intensive Care Unit; CDC, Centers for Disease Control and Prevention; NR, Not Reported; CI, Confidence Interval.

Table 3: Performance of the clinical risk scores for predicting SAP in the external validation cohorts

Score	External Validation cohort	Size (n)	Frequency of SAP (%)	SAP diagnostic criteria	Discrimination metric; C statistic (95% CI)	Calibration metric
ISAN	Athens	3204	12.8	Clinician reported	0.83 (0.81-0.85)	The Hosmer–Lemeshow goodness of fit test (Cox and Snell R^2 = 0.243)
A2DS2	NWGSR	45085	7.8	Clinician reported	0.83 (0.83-0.84)	Cox and Snell R^2 = 0.112, Nagelkerke R^2 =0.264, McFadden R^2 =0.215
	SSNAP	11648	6.7	Clinician reported	0.79 (0.77-0.81)	NR
	CNSR	8820	11.4	CDC	0.74 (0.73–0.75)	NR
	CNSR	5882	11.3	CDC	0.73 (0.72-0.74)	NR
	CICAS	3037	7.3	CDC	0.76 (0.74-0.77)	NR
	HNSR	1142	18.8	CDC	0.83 (0.8-0.87)	Cox and Snell R^2 = 0.243
	WCH	1279	24	Chinese Expert Consensus	NR	NR
	Shanghai	101	50.5	Chinese Expert Consensus	0.82(0.74-0.9)	NR
AIS-APS	CICAS	3 037	7.3	CDC	0.79 (0.76–0.82	The Hosmer–Lemeshow goodness of fit test; *P*=0.30
VHA score	CNSR	8820	11.4	CDC	0.75 (0.74–0.76)	NR
	CNSR	5882	11.3	CDC	0.73 (0.72–0.74)	NR
Pneumonia score	CNSR	8820	11.4	CDC	0.71 (0.70–0.72)	NR
	CNSR	5882	11.3	CDC	0.69 (0.68–0.71)	NR
	CICAS	3037	7.3	CDC	0.68 (0.66–0.69)	NR

SAP indicates Stroke-Associated Pneumonia; ISAN, Independence Prestroke, Sex, Age, National Institutes of Health Stroke Scale; A2DS2, Age, Atrial Fibrillation, Dysphagia, Sex, Severity; AIS-APS, Acute Ischaemic Stroke-Associated Pneumonia Score; VHA, Veterans Health Administration; NWGSR, North West Germany Stroke Register; SSNAP, Sentinel Stroke National Audit Programme; CNSR, Chinese National Stroke Registry; CICAS, Chinese Intracranial Atherosclerosis Study; HNSR, Henan Province Stroke Registry; WCH; Wuhan Central Hospital; CDC, Centers for Disease Control and Prevention; NR, Not Reported; CI, Confidence Interval.