Stroke

CLINICAL AND POPULATION SCIENCES

Course and Recognition of Poststroke Delirium

A Prospective Noninferiority Trial of Delirium Screening Tools

Robert Fleischmann, MD*; Sina Warwas*; Tina Andrasch; Rhina Kunz, MD; Carl Witt, MD; Annerose Mengel, MD; Bettina von Sarnowski, MD

BACKGROUND AND PURPOSE: Poststroke delirium (PSD) is an independent predictor of unfavorable outcome. Despite its individual and socioeconomic burden, its frequency, clinical course, and routine detection remain unresolved. This study aimed to assess psychometric properties of established delirium screening tools and investigate the natural course of PSD.

METHODS: This study investigated patients presenting with high-risk transient ischemic attacks or ischemic stroke within 24 hours during a 3-month period. Twice-daily screenings for PSD were done using the confusion assessment method, nursing delirium scale, and rapid delirium assessment, and evaluated for noninferiority against Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition criteria. We investigated demographic and stroke characteristics as predictors of PSD, neurological deficits as predictors of false screening results, and conducted a simulation study to estimate the best timing to identify PSD.

RESULTS: We enrolled 141 patients (73.8±10.4 years of age, 61 female) with a mean National Institutes of Health Stroke Scale score of 6.4±6.5. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition based PSD incidence was 39%, which manifested within 24 hours in 25% and 72 hours in almost all cases. The confusion assessment method was the only screening tool noninferior to Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition ratings providing a sensitivity of 82% and specificity of 80%. Age (odds ratio, 1.07 [1.02–1.13] per year, P=0.004) and National Institutes of Health Stroke Scale (odds ratio, 1.24 [1.15–1.34] per point, P<0.001) were predictors of PSD. False-positive screening results were associated with stroke-induced disorientation (odds ratio, 6.1 [3.2–11.61], P<0.001) and neglect (odds ratio, 2.17 [1.22–3.87], P=0.008). Simulations revealed that one in 4 cases is missed with less than daily screenings.

CONCLUSIONS: PSD is a common complication of stroke and transient ischemic attack. Detection is challenged by confounding effects such as focal neurological deficits and the necessity for at least daily screenings. Future studies are required to investigate implementation of these findings in clinical routine.

REGISTRATION: URL: https://www.clinicaltrials.gov. Unique identifier: NCT03930719.

Key Words: delirium ■ incidence ■ risk factors ■ screening ■ sensitivity and specificity ■ stroke

Despite advances of acute therapeutic options about 30% of patients with ischemic stroke remain with moderate, or worse, disability.[1] It has been recognized that specialized care of complications, such as dysphagia, venous thrombosis or embolism, and infections, is both viable and mandatory to further enhance the proportion of patients with satisfactory outcome.[2] Surprisingly, in this context poststroke delirium (PSD) has largely been neglected although it affects about 25% (range 10%–48%) of stroke patients,[3,4] and recent research indicates that an episode of delirium is not a fully reversible condition. In fact, it is an independent predictor of poor functional and cognitive outcome, increased complications rates, nursing times per patient, length of hospital stay, and long-term care dependency.[4] The growing elderly population is at an increased risk of

Correspondence to: Robert Fleischmann, MD, Department of Neurology, University Medicine Greifswald, Ferdinand-Sauerbruch-Str. 1, 17475 Greifswald, Germany. Email robert.fleischmann@uni-greifswald.de

*Dr Fleischmann and Ms Sina Warwas contributed equally.

For Sources of Funding and Disclosures, see page 477.

Stroke is available at www.ahajournals.org/journal/str

Nonstandard Abbreviations and Acronyms

4AT	rapid assessment test for delirium
CAM	confusion assessment method
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
IQCODE	Informant Questionnaire on Cognitive Decline in the Elderly
NIHSS	National Institutes of Health Stroke Scale
Nu-Desc	Nursing Delirium Screening Scale
OR	odds ratio
PSD	poststroke delirium
TIA	transient ischemic attack

PSD including its long-term negative effects.[5] Evidence for the prevention and care of PSD is hence urgently needed to meet current and imminent future challenges.[6]

See related article, p 479

Therefore, accurate, reliable, and easily applicable delirium screening tools and knowledge about the optimal frequency and timeline for their application are required. At present, there is insufficient evidence of how and when to screen for PSD. Only 4 well-designed studies evaluated the accuracy of delirium screening tools in stroke.[7] Three of these studies used the gold standard, that is, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), as comparator, but only one of these studies performed daily assessments acknowledging the fluctuating course of PSD.[8] The confusion assessment method (CAM) for the intensive care unit is a widely used modified version of CAM that includes operationalized instructions for the assessment of arousal and domains affected in delirium (attention, consciousness, etc) in critically ill, particularly ventilated, patients.[9] It was found to provide a sensitivity of 76% in patients with stroke but only 17% in another cohort.[10] The rapid assessment test for delirium (4AT) was reported to have a sensitivity of over 90% and specificity of 65% to 86%.[7] The Intensive Care Delirium Screening Checklist provided a high sensitivity of 98% but only moderate specificity of 55% in patients with aphasia.[8] The more commonly used CAM leads to very variable numbers of delirium (6.7%–42.6%) in several all stroke and hemorrhagic stroke populations that partly excluded patients with aphasia,[11] while the Nursing Delirium Screening Scale (Nu-Desc) has not been evaluated for PSD to date, although it provides high specificity and moderate sensitivity in twice-daily screenings on nursing units including one out of 4 neurological patients.[12]

This study aimed to fill this gap by assessing unselected patients with ischemic stroke for PSD through widely used delirium screening tools and comparing results with DSM-5 criteria. We were additionally interested in the natural course of PSD, which is another gap in the literature but provides critical information when to expect PSD and how screening frequency affects screening accuracy.

METHODS

Study Registration and Data Availability

This prospective randomized noninferiority trial was approved by the institutional ethics review board of the University Medicine Greifswald (BB 031/19) and adhered to the standards of the Helsinki declaration in its latest revision. The study was prospectively registered at https://www.clinicaltrials.gov. Source data are available from the corresponding author upon reasonable request.

Setting and Study Population

The study was conducted on an Acute Stroke Unit of the University hospital Greifswald for a continuous 3-month period starting in May 2019, including weekends and public holidays. All patients presenting within 24 hours after symptom onset with ischemic stroke or high-risk transient ischemic attack (TIA) defined by an ABCD2 score (ie, a sum of scores given for age, blood pressure, clinical features, duration of symptoms and diabetes that are each given 0–2 points depending on their presence and severity) of ≥6 and according to the world health organization's definition were eligible for enrollment. If the onset of stroke was unclear, the time when the patient was last seen well served as onset. Patients with hemorrhagic stroke were excluded. There were no further exclusion criteria, such as severe aphasia, to enhance generalizability of our findings and since DSM-5 criteria do not necessitate verbal communication, but presentation of visual or acoustic stimuli is sufficient.[3,11] All patients, or their legal representative, provided written informed consent. Data collection, but not evaluation, was allowed to begin immediately after admission.

Study Design, Delirium Assessment, and Primary Outcome Measures

The study was carefully designed to enable state-of-the art assessments and comparisons of PSD detection rates using screening tools and DSM criteria as primary outcome measures. Patient evaluation was conducted by delegated personnel to avoid possible interference with routine treatment by repeated assessments through several screening tools in the study setting that would not have occurred in clinical routine. Furthermore, this approach allowed control of interrater variability. All patients required a sufficient level of arousal for further testing measured with the Richmond Agitation-Sedation Scale using a cutoff of ≥−3, that is, moderate sedation or better. Clinical assessment for PSD was done twice daily, that is, during morning (8 AM–3 PM) and late shifts (3 PM–10 PM) with an interval of about 8 hours. Assessments were standardized

 Stroke. 2021;52:471–478. DOI: 10.1161/STROKEAHA.120.031019

and performed by one of 2 medical students trained in application of delirium screening tools (S. Warwas and T. Andrasch), and one of 2 board-certified neurologists using DSM-5 criteria (Drs Witt and Fleischmann). Students and neurologists aimed to conduct their investigations within a comparable time frame (ie, within the beginning or end of a shift but there were no mandatory intervals) and were blinded to screening and DSM test results, respectively. Chart reviews were performed every morning to evaluate night shifts (10 PM to 8 AM), which has been validated to complement on-site delirium screening.[13] PSD assessment lasted for a maximum of 7 days following symptom onset or until discharge or death of the patient. PSD screening tools were chosen based on their absence of copyright restrictions and thus broad applicability, and specifically given their popularity (CAM), quick and easy administration through nurses (Nu-Desc), and prior use in stroke studies (4AT).[14,15] They were applied in randomized order to avoid sequence effects.

If a patient was tested positive for PSD according the DSM-5 criteria, treating physicians were informed about the result and followed national guidelines for the general management of delirium since no guidelines specifically designed for PSD were available.[16]

Secondary Outcome Parameters and Predictors of Delirium Screening Accuracy

Stroke subtypes were classified according to the Oxfordshire Community Stroke Project, that is, as partial anterior circulation stroke, total anterior circulation stroke, posterior circulation stroke, or lacunar infarction.[17] Stroke cause was defined by Trial of ORG 10172 in Acute Stroke Treatment criteria.[18]

The National Institutes of Health Stroke Scale (NIHSS) was concomitantly assessed with every PSD evaluation in the first 72 hours since accuracy of delirium screening tools may be flawed by focal neurological impairment.[2,7] Patients' prestroke cognitive abilities potentially predispose for PSD and were assessed through the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE).[19] Finally, we investigated the delirium screening frequency required not to miss any of the PSD cases through a simulation based on this study's data.

Data Management, Sample Size Considerations, and Statistics

Statistical evaluations were done using SPSS (Version 25, IBM, Armonk, NY). Simulations were run using Matlab (R2018a, Natick, MA). Results from descriptive statistics are reported as mean ± SD or median and interquartile range depending on data distribution. Results from inferential statistics are reported with their appropriate coefficients and, if applicable, odds ratios (OR) including 95% CI in brackets and *P* values denoting the statistical significance.

The primary hypothesis, that is, that the accuracy of screening tools was noninferior to DSM-5 results, was tested using the McNemar test that accounts for paired binary observation of PSD frequencies. For this purpose, test results of each detection method were summarized to a second level binary response variable characterizing a patient as PSD positive or negative at any point during the study period. In other words, a patient with PSD on day 1 and 5 would score only once as delirious. This approach was chosen to reduce alpha error inflation and since our primary hypothesis was the recognition of PSD per se, which is considered most important for clinical routine management and patient outcome.[6,20] In a 3-month period, we expected to include about 200 patients leading to the detection of small differences (effect size $w \geq 0.25$) with a 2-sided α-error of 5% and power of 90%. Agreement of screening tools with DSM-5 results is additionally reported through Cohen κ coefficients. Receiver operating curve statistics were used to calculate the area under the curve, sensitivity, and specificity of screening tools. Interrater reliabilities, that is, agreement between the 2 raters using screening tools or DSM criteria, respectively, were assessed using intraclass correlations with 1-way random effects and included their 95% CI.

Multiple binomial logistic regression was used to evaluate demographic and stroke characteristics as predictors of PSD. The full model included sex, age, NIHSS peak score, and Oxfordshire Community Stroke Project type. NIHSS item scores, NIHSS total score, and IQCODE scores were additionally investigated as predictors of screening tool accuracy. High-frequency study data were sampled down to simulate the rate of missed PSD cases if screening was done on (1) just 1, (2) 2 or (3) 3 random occasions during the complete inpatient period, (4) once daily at a random time, or (5) once daily during morning or (6) late shifts. Simulations were run 100 000 times per situation, and bootstrapping was used to estimate 95% CI. *P* values equal to or lower than 0.05 are considered significant, values lower than 0.001 are not reported exact but as <0.001. Multiple comparisons are corrected for α-error accumulation by a Bonferroni correction.

RESULTS

During the study period, 327 patients were admitted to the stroke unit and screened for eligibility. There were 115 patients with ischemic stroke or TIA that did either not present within 24 hours of stroke onset or did not have a high-risk TIA, 16 patients presented with intracranial hemorrhage, and 47 with stroke mimics. Six eligible patients did not consent to participate and 2 patients withdrew their consent, hence 141 patients (73.8±10.4 years of age, 61 female) were enrolled and completed the study. Seven patients had a high-risk TIA. Mean NIHSS score at admission was 6.4±6.5. Clinical stroke localization was partial anterior circulation stroke in 59% (n=79), total anterior circulation stroke in 7.5% (n=10), lacunar infarction in 14.9% (n=20), and posterior circulation stroke in 18.7% (n=25). Stroke causes were large artery atherosclerosis in 17% (n=24), cardioembolism in 36.9% (n=52), small-vessel occlusion in 16.3% (n=23), other causes in 7.8% (n=11) cases, and undetermined in 22% (n=31). The median observation period was 6 days (interquartile range, 3–6). A summary of the population characteristics is given in Table 1. There was no missing data unless reported otherwise.

PSD Incidence, Time Course, and Predictors

Thirty-nine percent of patients with stroke (n=55) developed an episode of delirium based on expert ratings

Table 1. Tabular View of Population and Stroke Characteristics

	All patients	With delirium	Without delirium	Statistical difference*
Age, y	73.8±10.4	75.8±13.8	71.6±10.5	0.041†
Sex (male/female)	80/61	33/22	47/39	0.532
NIHSS (on admission)	6.4±6.5	15±8.4	5.6±5.2	<0.001†
IQCODE	3.07±0.53	3.21±0.45	3.02±0.55	0.224
Stroke subtypes				0.043†
PACI	79	32	47	0.899
TACI	10	7	3	0.051
LACI	20	4	16	0.112
POCI	25	10	15	Ref
Stroke cause, n				0.285
Large artery atherosclerosis	24	9	15	…
Cardioembolism	52	24	28	…
Small-vessel occlusion	23	8	15	…
Other causes	11	2	9	…
Undetermined	31	12	19	…

Patients with an episode of delirium were older and more severely affected but did not differ with regards to prestroke cognitive functioning or gender. Stroke cause was not a significant predictor of an episode of delirium; hence no post hoc testing was performed. Stroke subtype yielded a significant global effect, but TACI was only marginally significant to be more often delirious as compared to nondelirious patients. IQCODE indicates Informant Questionnaire on Cognitive Decline in the Elderly; LACI, lacunar infarction; NIHSS, National Institutes of Health Stroke Scale; PACI, partial anterior circulation stroke; POCI, posterior circulation stroke; ref, reference category in binary logistic regression; and TACI, total anterior circulation stroke.

*Two-tailed *t* test for unpaired samples for continuous data, χ^2 test for 2-level nominal data, binary logistic regression for higher-level nominal data.

†Difference between group statistically significant at an alpha level of 0.05.

using DSM-5 criteria. Chart review of night shifts did not add any new cases, thus only on-site screenings were considered for further evaluations. Motor subtypes were hyperactive in 7% (n=4), mixed in 58% (n=32), and hypoactive in 35% (n=19). Ten percent of patients with stroke (n=14) developed PSD within 24 hours following stroke onset and almost all cases (98%, n=54) occurred within 72 hours (Figure). Mean duration of PSD was 48±42 hours. Gender was not a significant predictor for PSD (*P*=0.23). Stroke location was a significant global factor (*P*=0.04) but post hoc analyses revealed only a nonsignificant trend for total anterior circulation stroke (OR, 0.99 [0.93–122.11], *P*=0.051). Age (OR, 1.07 [1.02–1.13] per year, *P*=0.004) and peak NIHSS (OR, 1.24 [1.15–1.34] per point, *P*<0.001) were clear predictors of PSD. Proxies of all patients were handed out an IQCODE scoring sheet, yet only n=62 (44%) were returned and thus available for analysis. Based on this data set, IQCODE was also not a significant predictor of PSD (*P*<0.22). There was no significant correlation between predictors.

Accuracy of PSD Screening Tools, Its Predictors, and Interrater Reliability

PSD frequency observed with the CAM (n=58) did not significantly differ from the gold standard (*P*=0.25), the κ coefficient for agreement was 0.61. The 4AT (n=65) and Nu-Desc (n=77) identified statistically significant more PSD patients than the gold standard (*P*=0.016 and <0.001, respectively), their κ coefficients were 0.57 and 0.55, respectively. Receiver operating curve analyses revealed an area under the curve of 0.81 for the CAM, 0.80 for the 4AT, and 0.79 for the Nu-Desc. Details of their diagnostic performance, including sensitivity, specificity, and predictive values, are given in Table 2.

The prestroke IQCODE did not predict any false-negative or -positive screening results. The NIHSS sum score was a significant predictor of false-negative CAM screening results (OR, 1.12 [1.03–1.23], *P*=0.013) but did not define any other false-negative or -positive screening results. None of the single NIHSS items significantly predicted false-negative screening results. However, false-positive screening results were affected by several NIHSS items as follows. The CAM false positively identified patients with minor disorientation (OR, 6.1 [3.2–11.61], *P*<0.001) and minor neglect (OR, 2.17 [1.22–3.87], *P*=0.008) as being delirious. Major disorientation reduced the odds of being false positively classified delirious (OR, 0.51 [0.28–0.92], *P*=0.025). The 4AT and Nu-Desc false positively identified patients with minor disorientation (OR_{4AT}, 4.02 [2.3–7.01], *P*<0.001 and $OR_{Nu\text{-}Desc}$, 2.05 [1.22–3.44], *P*=0.007) and minor language disturbance (OR_{4AT}, 2.11 [1.05–4.25], *P*=0.035 and $OR_{Nu\text{-}Desc}$, 2.61 [1.19–5.7], *P*=0.017). Minor disturbances following commands significantly reduced the odds of being false positively classified delirious by the Nu-Desc (OR, 0.32 [0.13–0.79], *P*=0.013).

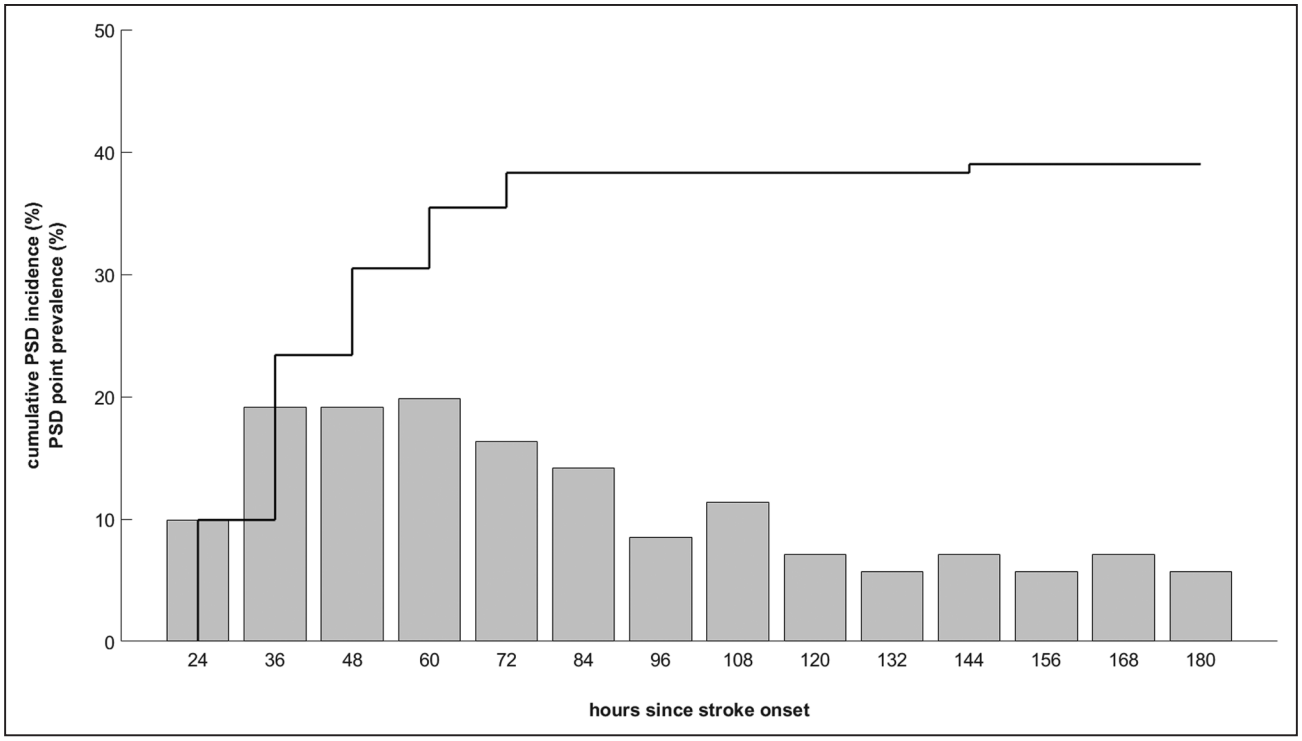

Figure. Poststroke delirium (PSD) incidence and point prevalence.
Cumulative incidence and point prevalence of stroke shows a steep increase in new PSD cases within the first 72 h of stroke onset and only one case of later PSD at 144 h. PSD prevalence was highest between 36 and 60 h following stroke onset indicating that delirium was transient. The fluctuating nature of PSD is indicated by changes of its prevalence, although no new cases were added between 84 and 132 h.

Intraclass correlation coefficients indicated good interrater agreement and were 0.65 (0.62–0.68) for DSM-5, 0.66 (0.63–0.69) for CAM, 0.64 (0.61–0.67) for 4AT, and 0.61 (0.57–0.64) for Nu-Desc ratings.

Simulation of Less Frequent PSD Screening and Potentially Missed PSD Cases

The simulation revealed that 49.2% (38.7%–58.2%) of PSD cases would have been missed if screening had been done only once, 34.1% (24.9%–42.9%) if it had been done twice and 25.3% (17.8%–33.3%) if it had been done 3 times within the screening period. Daily screenings at a random time would have caused a rate of missed PSD cases of 5.7% (1.8%–14.7%) while daily screenings during morning or late shifts would have caused a missing rate of 7.3% (1.8%–16.4%) or 3.6% (0%–12.1%), respectively.

DISCUSSION

Our study shows that patients with ischemic stroke or high-risk TIA treated on a dedicated stroke unit have a high risk of developing delirium within the first 72 hours after symptom onset. The CAM was the best screening tool to identify PSD and the chance of

Table 2. Diagnostic Performance of Delirium Screening Tools in Poststroke Delirium

Tool	No. of patients with PSD	% of patients with PSD	Statistical significance vs DSM-5 rating	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Confusion assessment method	n=58	41.1%	*P*=0.248	82% (71%–91%)	80% (70%–87%)	72% (64%–80%)	87% (80%–92%)
Rapid assessment test for delirium	n=65	46.1%	*P*=0.016	86% (73%–93%)	74% (63%–82%)	68% (60%–76%)	89% (82%–93%)
Nursing delirium screening scale	n=77	54.6%	*P*<0.001	93% (83%–98%)	66% (56%–76%)	64% (55%–72%)	94% (88%–97%)

Trained neurologists using DSM-5 criteria, ie, the international gold standard for the diagnosis of delirium, identified n=55 (39%) patients with an episode of PSD. The confusion assessment method was the only tool that identified an equal frequency of PSD cases that did not statistically differ from DSM-5 ratings in a McNemar test that accounts for paired observations. The confusion assessment method furthermore provided a diagnostic performance with sensitivity, specificity, and predictive values of about 80%, while comparators were generally more sensitive but less specific or predictive for an episode of PSD. Bootstrap 95% CIs of sensitivity, specificity, and predictive values are given in brackets. DSM-5 indicates Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; and PSD, poststroke delirium.

detection was especially high when daily screenings, particularly during late shifts, are performed. Screening results were affected by acute but not by prestroke cognitive dysfunction.

PSD Incidence in Relation to Previous Studies and Potential Effects of Screening Frequency

The incidence of PSD in this study was surprisingly high compared with previous studies. The most recent and comprehensive review on the occurrence rate of PSD by Shaw et al[11,21] concluded that an incidence of about 25% [20%–30%] is to be expected, which they confirmed in a prospective study, and only those studies including intracerebral hemorrhages revealed higher rates up to 48%. Yet, the review identified a possibly serious risk of bias owing to patient selection since patients at the highest risk, such as those with dementia or severe stroke, were partly excluded. The quality of evidence of the meta-analysis was concluded to be only moderate. It is hence plausible that the lack of major exclusion criteria yielded higher delirium rates in our study. Our patients were furthermore older and more severely affected as compared to the study by Shaw et al,[11] which both increase the risk for PSD. The predominance of hypoactive and mixed motor subtypes with rare hyperactive cases provides another plausible explanation for lower PSD rates in previous studies since hypoactive variants often remain unrecognized.[20] It is also possible that previous studies missed cases of PSD since they included subacute strokes up to 7 days following symptom onset.[11] Simulation results of less frequent PSD screenings yielded another possible explanation of underreported PSD rates and revealed that daily screenings are required not to miss an episode of delirium. This supports the application of current guidelines for the detection of intensive care delirium for stroke patients since >25% of cases are potentially missed when screening is not done at least on a daily basis.[22] Simulation results, however, do not clearly show that multiple screenings per day are required since the lower bound of the confidence interval of late shift screenings reached a missing rate of 0%. This may either be an effect of sample size, but one would also expect that delirium rather occurs in later hours given its circadian rhythmicity.[23] In line with this notion, the addition of chart-based screening for delirium during night shifts can increase the diagnostic yield by about 50%, which might be a viable option in PSD but requires validation since chart review did not add any cases in our study.[13]

Performance of Screening Tools for the Detection of PSD

The CAM was the only screening tool that was noninferior to the gold standard. Its substantial agreement with DSM-5 results, and sensitivity and specificity of about 80%, render it a suitable screening tool for the detection of PSD. Its sensitivity is lower than that reported for the Intensive Care Delirium Screening Checklist, it is yet more specific and the Intensive Care Delirium Screening Checklist was validated in a group of patients two-thirds of which suffered from aphasia rendering direct comparisons difficult.[8] The 4AT and Nu-Desc yielded PSD rates that significantly differed from the gold standard and their agreement with DSM-5 results was only moderate. Their use can, therefore, not be endorsed based on this study's result. This recommendation may be in apparent contradistinction to previous findings, yet none of the previous studies directly tested for agreements of methods but rather based their interpretations on test sensitivity and specificity that are not suitable statistical comparisons of diagnostic tests.[7,24] While the lack of major exclusion criteria enhances generalizability of our results, they require replication in an independent sample. It must furthermore be examined whether the application of the CAM is feasible and its diagnostic performance applicable in clinical routine or possibly confounded, for example, since it is more difficult to apply than the Nu-Desc or 4AT.[25] Moreover, it must be duly noted that screening tools critically depend on the patient's cognitive abilities and that in particular the CAM was designed for use in conjunction with a separate cognitive test, whereas the 4AT has very brief tests of cognition within it.[15,26] This necessity may influence the applicability of screening tools in clinical routine, yet our results confirm that, even without cognitive testing, the CAM is accurate in detecting PSD. Another concern is that screening for PSD at one point in time may miss changes that fluctuate. While chart review did not reveal any additional cases of PSD in this study, previous research highlighted the utility of board rounds to enhance the detection rate of delirium.[27] Future studies need to validate this approach in PSD since our results clearly show that neurological deficits impact on patient's cognitive abilities and lead to erroneous screening results.

Predictors of Screening Tool Performance

False-negative screening results were only affected by higher NIHSS sum scores, which indicates that disentanglement of stroke and delirious features is difficult. However, there were no single predictors of false-negative screening results. Nonetheless, the sensitivity of the best performing tool, that is, the CAM, was lower than 94% [91%–97%] as reported in a comprehensive review in other populations.[28] It is recognized that the CAM may be difficult to apply by untrained personnel, which provides a plausible explanation despite our efforts to provide sufficient familiarity with the tools.[29] Supporting this view, the Nu-Desc had a 10% higher sensitivity for the detection of PSD and is known to be easiest to apply.[30] Future studies with longitudinal

 Stroke. 2021;52:471–478. DOI: 10.1161/STROKEAHA.120.031019

assessments could be used to evaluate training effects on PSD screening accuracy.

The specificity of all tools was clearly confounded by particular stroke features. Unsurprisingly, disorientation as a direct consequence of stroke, for example, due to lesions of the precuneus or prefrontal structures, led to misclassifications of patients as being delirious irrespective of the tool used.[31] Other predictors of false-positive rates affected only single tools but also comprised of higher-order cognitive disorders, such as inability to follow commands, neglect, and aphasia. Interestingly, only mild symptoms were confused as features of PSD while severe inability to follow commands, global aphasia or multimodal neglect were not significantly associated with false-positive rates. This highlights that PSD is a challenging entity that requires recognition of minor variants of focal neurological deficits and strict adherence to DSM-5 criteria that not only demand deficient cognition and awareness but also inattention as a core feature.[3] Stroke specific adaptations of screening tools might be beneficial to enhance their accuracy in the presence of neurological deficits as was previously done with the intensive care adaptation of the CAM, that is, CAM-ICU, to overcome challenges due to mechanical ventilation.[9]

It may seem surprising that prestroke cognitive impairment quantified by the IQCODE did not confound test accuracy since underdiagnosis of delirium is common in dementia.[20] Meagher et al,[32] however, found that impaired attention, being a key feature of delirium, is an important distinguishing feature against dementia and screening for a disorder of attention is part of all screening tools used in this study. Importantly, we trained study personnel to be particularly accurate when testing for impaired attention in anticipation of coexistence of dementia and delirium. However, <50% of IQCODE proxy ratings were available, and results thus not powered to be confident that prestroke cognitive impairment is not a significant predictor of false-positive screening results. Particularly given to the poor specificity of the 4AT future studies are required to confirm or refute that possibility.

Limitations

The study recruited fewer patients than expected, which may have affected our ability to detect differences between CAM and DSM-5 ratings, which we consider equivalent based on our study results. Post hoc analyses of the actual power for effect sizes of $w \geq 0.25$ yielded 84% as compared to 90% that were initially planned. Since a power of ≥80% is generally considered acceptable, we argue that our results should not be substantially biased by sample size and hence be applicable to a general population of patients with stroke. Yet, the lower than expected sample size inevitably confounds the examination of predictors of PSD, for example, the influence of stroke subtypes and prestroke cognitive abilities. This includes the IQCODE that is additionally confounded by a low return rate of scoring sheets and renders its analysis preliminary. Another limitation is that this study was monocentric, which clearly affects generalizability. Yet, the study design was chosen to avoid selection bias and include an unselected consecutive cohort of patients, which should, in turn, enhance generalizability. The next step would be to perform a multicenter study that also includes patients with hemorrhagic stroke since it would be impractical to use different tools in different subsets of patients. Our study results are furthermore possibly confounded by uncertainties regarding the onset of stroke in 17% of cases (n=21). Since this proportion is typical in patients with stroke, we are unaware of means to improve this inaccuracy. It is possible that the true course of PSD may differ from our estimation, but we expect only minor deviations since patients with possible onset beyond 24 hours were excluded, and hence our results should nonetheless help plan future studies investigating pathomechanisms and interventions.

Conclusions

This study found that PSD affects about 40% of patients with acute stroke or high-risk TIA. PSD typically develops within 72 hours following symptom onset and one in 4 cases occurs even within the first 24 hours. Detection of PSD is particularly challenging since features of stroke can be confused with delirium. Consequently, the CAM was the only tool studied that was noninferior to expert ratings using DSM-5 criteria. At least one in 4 PSD cases remains undetected if screenings are done on a less than daily basis.

ARTICLE INFORMATION

Received May 24, 2020; final revision received October 20, 2020; accepted November 9, 2020.

Affiliations

Department of Neurology, University Medicine Greifswald, Germany (R.F., S.W., T.A., R.K., C.W., B.v.S.). Department of Neurology and Stroke, Hertie-Institute for Clinical Brain Research, Eberhard-Karls University of Tübingen, Germany (A.M.).

Sources of Funding

None.

Disclosures

None.

REFERENCES

1. Kunz A, Ebinger M, Geisler F, Rozanski M, Waldschmidt C, Weber JE, Wendt M, Winter B, Zieschang K, Fiebach JB, et al. Functional outcomes of pre-hospital thrombolysis in a mobile stroke treatment unit compared with conventional care: an observational registry study. *Lancet Neurol.* 2016;15:1035–1043. doi: 10.1016/S1474-4422(16)30129-6
2. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, Biller J, Brown M, Demaerschalk BM, Hoh B, et al; American Heart Association Stroke Council. 2018 guidelines for the early management of

patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* 2018;49:e46–e110. doi: 10.1161/STR.0000000000000158
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, (dsm-5®).* American Psychiatric Pub; 2013.
4. Shi Q, Presutti R, Selchen D, Saposnik G. Delirium in acute stroke: a systematic review and meta-analysis. *Stroke.* 2012;43:645–649. doi: 10.1161/STROKEAHA.111.643726
5. He W, Goodkind D, Kowal P. An aging world: 2015. *International Population Reports.* U.S. Census Bureau; 2016.
6. Qu J, Chen Y, Luo G, Zhong H, Xiao W, Yin H. Delirium in the acute phase of ischemic stroke: incidence, risk factors, and effects on functional outcome. *J Stroke Cerebrovasc Dis.* 2018;27:2641–2647. doi: 10.1016/j.jstrokecerebrovasdis.2018.05.034
7. Mansutti I, Saiani L, Palese A. Detecting delirium in patients with acute stroke: a systematic review of test accuracy. *BMC Neurol.* 2019;19:310. doi: 10.1186/s12883-019-1547-4
8. Boßelmann C, Zurloh J, Stefanou MI, Stadler V, Weber Y, Lerche H, Poli S, Ziemann U, Mengel A. Delirium screening in aphasic patients with the intensive care delirium screening checklist (ICDSC): a Prospective Cohort Study. *Front Neurol.* 2019;10:1198. doi: 10.3389/fneur.2019.01198
9. Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, Truman B, Speroff T, Gautam S, Margolin R, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA.* 2001;286:2703–2710. doi: 10.1001/jama.286.21.2703
10. Serpa Neto A, Slooter AJ. Delirium detection in stroke patients. *Crit Care Med.* 2012;40:2266–2267; author reply 2267. doi: 10.1097/CCM.0b013e31824fc115
11. Shaw RC, Walker G, Elliott E, Quinn TJ. Occurrence rate of delirium in acute stroke settings: systematic review and meta-analysis. *Stroke.* 2019;50:3028–3036. doi: 10.1161/STROKEAHA.119.025015
12. Hargrave A, Bastiaens J, Bourgeois JA, Neuhaus J, Josephson SA, Chinn J, Lee M, Leung J, Douglas V. Validation of a nurse-based delirium-screening tool for hospitalized patients. *Psychosomatics.* 2017;58:594–603. doi: 10.1016/j.psym.2017.05.005
13. Saczynski JS, Kosar CM, Xu G, Puelle MR, Schmitt E, Jones RN, Marcantonio ER, Wong B, Isaza I, Inouye SK. A tale of two methods: chart and interview methods for identifying delirium. *J Am Geriatr Soc.* 2014;62:518–524. doi: 10.1111/jgs.12684
14. Gaudreau JD, Gagnon P, Harel F, Tremblay A, Roy MA. Fast, systematic, and continuous delirium assessment in hospitalized patients: the nursing delirium screening scale. *J Pain Symptom Manage.* 2005;29:368–375. doi: 10.1016/j.jpainsymman.2004.07.009
15. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med.* 1990;113:941–948. doi: 10.7326/0003-4819-113-12-941
16. Martin J, Heymann A, Bäsell K, Baron R, Biniek R, Bürkle H, Dall P, Dictus C, Eggers V, Eichler I, et al. Evidence and consensus-based German guidelines for the management of analgesia, sedation and delirium in intensive care–short version. *Ger Med Sci.* 2010;8:Doc02. doi: 10.3205/000091
17. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet.* 1991;337:1521–1526. doi: 10.1016/0140-6736(91)93206-o
18. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE III. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke.* 1993;24:35–41. doi: 10.1161/01.str.24.1.35
19. Jorm AF, Jacomb PA. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, validity and some norms. *Psychol Med.* 1989;19:1015–1022. doi: 10.1017/s0033291700005742
20. Ritter SRF, Cardoso AF, Lins MMP, Zoccoli TLV, Freitas MPD, Camargos EF. Underdiagnosis of delirium in the elderly in acute care hospital settings: lessons not learned. *Psychogeriatrics.* 2018;18:268–275. doi: 10.1111/psyg.12324
21. Shaw R, Drozdowska B, Taylor-Rowan M, Elliott E, Cuthbertson G, Stott DJ, Quinn TJ. Delirium in an acute stroke setting, occurrence, and risk factors. *Stroke.* 2019;50:3265–3268. doi: 10.1161/STROKEAHA.119.025993
22. Aldecoa C, Bettelli G, Bilotta F, Sanders RD, Audisio R, Borozdina A, Cherubini A, Jones C, Kehlet H, MacLullich A, et al. European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium. *Eur J Anaesthesiol.* 2017;34:192–214. doi: 10.1097/EJA.0000000000000594
23. Maldonado JR. Delirium pathophysiology: an updated hypothesis of the etiology of acute brain failure. *Int J Geriatr Psychiatry.* 2018;33:1428–1457. doi: 10.1002/gps.4823
24. Hawass NE. Comparing the sensitivities and specificities of two diagnostic procedures performed on the same group of patients. *Br J Radiol.* 1997;70:360–366. doi: 10.1259/bjr.70.832.9166071
25. De J, Wand AP. Delirium screening: a systematic review of delirium screening tools in hospitalized patients. *Gerontologist.* 2015;55:1079–1099. doi: 10.1093/geront/gnv100
26. Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ. Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in an acute stroke unit setting. *Stroke.* 2013;44:3078–3083. doi: 10.1161/STROKEAHA.113.001724
27. Pendlebury ST, Lovett NG, Smith SC, Dutta N, Bendon C, Lloyd-Lavery A, Mehta Z, Rothwell PM. Observational, longitudinal study of delirium in consecutive unselected acute medical admissions: age-specific rates and associated factors, mortality and re-admission. *BMJ Open.* 2015;5:e007808. doi: 10.1136/bmjopen-2015-007808
28. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The confusion assessment method: a systematic review of current usage. *J Am Geriatr Soc.* 2008;56:823–830. doi: 10.1111/j.1532-5415.2008.01674.x
29. Wong EK, Lee JY, Surendran AS, Nair K, Della Maestra N, Migliarini M, St Onge JA, Patterson CJ. Nursing perspectives on the confusion assessment method: a qualitative focus group study. *Age Ageing.* 2018;47:880–886. doi: 10.1093/ageing/afy107
30. van Velthuijsen EL, Zwakhalen SM, Warnier RM, Mulder WJ, Verhey FR, Kempen GI. Psychometric properties and feasibility of instruments for the detection of delirium in older hospitalized patients: a systematic review. *Int J Geriatr Psychiatry.* 2016;31:974–989. doi: 10.1002/gps.4441
31. Peer M, Salomon R, Goldberg I, Blanke O, Arzy S. Brain system for mental orientation in space, time, and person. *Proc Natl Acad Sci U S A.* 2015;112:11072–11077. doi: 10.1073/pnas.1504242112
32. Meagher DJ, Leonard M, Donnelly S, Conroy M, Saunders J, Trzepacz PT. A comparison of neuropsychiatric and cognitive profiles in delirium, dementia, comorbid delirium-dementia and cognitively intact controls. *J Neurol Neurosurg Psychiatry.* 2010;81:876–881. doi: 10.1136/jnnp.2009.200956