

RESEARCH

Open Access



Delirium detection in hospitalized adults: the performance of the 4 'A's Test and the modified Confusion Assessment Method for the Emergency Department. A comparison study

Olga Muser^{1,2*} , Kevin Seiler³, Stefanie Bachnick^{1,4}, Beatrice Gehri^{1,5}, Franziska Zúñiga¹ and Wolfgang Hasemann^{3,6}

Abstract

Background: Early detection of delirium through systematic screening is essential to mitigate and prevent possible consequences. The 4 'A's Test (4AT) is a new tool that can be used for delirium detection easily and without special training. The modified Confusion Assessment Method for Emergency Department (mCAM-ED) is an operationalized version of the Confusion Assessment Method, a worldwide used tool for delirium screening in clinical practice and research. This is the first comparison of both delirium screening tools. This study aimed to investigate performance accuracy of the 4AT compared to the mCAM-ED in detecting delirium in hospitalized patients.

Methods: In this prospective single-centre cross-sectional pilot study, patients from six wards were selected consecutively. All patients underwent a delirium screening with the gold standard, the mCAM-ED. To rate the algorithm of the 4AT, corresponding items of the mCAM-ED were derived and used.

Results: A total of 116 patients with a median age of 73 years could be included. Dementia was present in 11 (9.5%) patients, and 42.2% were women. Delirium was present in 8/116 (6.9%) and 16/116 (13.8%) patients according to the mCAM-ED and the 4AT, respectively. In comparison, the 4AT showed 100% (95% CI 0.63, 1.00) sensitivity, 93% (95% CI 0.86, 0.97) specificity, 13.50 (95% CI 6.93, 26.30) positive likelihood ratio and 0.00 (95% CI 0.00, NaN) negative likelihood ratio.

Conclusions: In this first comparison, the 4AT showed a high rate of false-positive scores, which may result in an increased need for further in-depth assessments.

Keywords: Delirium, CAM, mCAM-ED, 4AT, Performance comparison, Sensitivity and specificity

Background

Delirium is an acute cognitive disorder with disturbances in attention, awareness and further cognitive functions. It develops in a short period of time, and the severity of its symptoms fluctuates throughout the day. The disorder is a response to one or more pathophysiological factors due to a medical condition, substance intoxication,

*Correspondence: olga@muser.cc

¹ Department Public Health, Faculty of Medicine, Institute of Nursing Science, University of Basel, Bernoullistrasse 28, 4056 Basel, Switzerland
Full list of author information is available at the end of the article

withdrawal or exposure to a toxin (American Psychiatric Association 2013).

Delirium occurs in people of every age (Inouye et al. 2014) and people who are older than 65 years have a higher risk of developing delirium (Inouye 2006). In the Western world, prevalence of delirium on general medical wards is between 18–35% and 17% on surgical wards (Inouye et al. 2014). Delirium is associated with several deteriorating consequences for patients, health professionals and health care systems (Inouye et al. 2014; Leslie et al. 2008; Wand et al. 2013). Individuals with delirium show a higher mortality rate after five years (Moskowitz et al. 2017). Delirium can lead to progression of existing dementia (Davis et al. 2017) and to an increased risk of developing dementia (Olofsson et al. 2018). Each additional day of delirium persistence decreases self-efficacy in terms of daily activities and cognition (Han et al. 2017). Delirium is also a very emotionally stressful situation for patients and their relatives (Cohen et al. 2009). The increased care needs of patients with delirium place enormous demands on nurses in this situation (Belanger and Ducharme 2011). In addition, patients suffering from delirium have 2.5 times the average daily health care costs in a hospital compared to patients without delirium (Leslie et al. 2008).

Early detection of delirium is crucial to mitigate and prevent its potential consequences (Mittal et al. 2011) and can be achieved with systematic screening (Inouye et al. 2001; Grossmann et al. 2014). While early and precise detection of delirium enables early intervention and treatment of the underlying cause (Inouye 2006; Inouye et al. 2014), up to 80% of delirium is undetected by healthcare professionals (Inouye et al. 2001; Collins et al. 2010). Additionally, causal treatment is thus often delayed or missed. Therefore, it is vital that the screening tools perform well, and are time efficient (De and Wand 2015).

As the Diagnostic and Statistical Manual of Mental Disorders (DSM) has evolved, new screening tools have been developed or existing screening tools have been enhanced (Fick et al. 2015), such as the 4 'A's Test (4AT) (MacLulich et al. 2011) and the modified Confusion Assessment Method for the Emergency Department (mCAM-ED) (Grossmann et al. 2014; Hasemann et al. 2018a). The 4AT is a concise and easy-to-use tool (De et al. 2017) that does not require special training (Bellelli et al. 2014). The mCAM-ED is a modified version of the Confusion Assessment Method (CAM) (Inouye et al. 1990), a tool for delirium screening in clinical practice and research that is used worldwide (Wei et al. 2008).

Both these tools detect and recognize delirium in the same cognitive domains (i.e. consciousness, attention and cognition). The difference between the tools is that the

4AT uses fewer questions and does not assess disorganized thinking. The mCAM-ED requires training and is therefore more complex to implement (Grossmann et al. 2014). Furthermore, when the CAM was used without a formal assessment, nurses recognized that delirium was present in only 19% of cases, leading to a misinterpretation rate of delirium of up to 80% (Inouye et al. 2001). The mCAM-ED solved this disadvantage by defining the items (Grossmann et al. 2014; Hasemann et al. 2018a). As the 4AT is less complex, a comparison of the 4AT performance in delirium detection with the mCAM-ED is of great interest. However, to our knowledge, no comparison between the 4AT and the mCAM-ED has yet been performed. The aim of this study was to investigate the accuracy of the performance, i.e. sensitivity and specificity, positive and negative likelihood, and positive and negative predictive values, of the 4AT compared to the mCAM-ED in detecting delirium in hospitalized patients. The expectation when comparing the two tools was that they would produce relatively similar results, with the 4ATs showing a sensitivity and specificity of 95% each. We conducted this small preliminary study as a pilot study to evaluate the feasibility to rate the 4AT based on the mCAM-ED assessment.

Methods

Design

This study is a prospective single-centre cross-sectional pilot study, which investigated the performance accuracy of the 4AT. The mCAM-ED was applied as a gold standard.

Setting and sample

This pilot study was carried out in the University Hospital Basel, a tertiary care centre in north-western Switzerland. In 2018, about 38,000 patients were hospitalized at the University Hospital Basel. Overall, the University Hospital Basel has a total capacity of 770 beds (Universitätsspital Basel 2019). Patients were consecutively selected from six wards with a total capacity of approximately 200 beds. The following disciplines were included: orthopaedics, traumatology and spinal disorders, cardiovascular, pulmonary and metabolic diseases, diseases of the blood and blood-forming organs, kidney diseases and tumours of various organ systems.

Patients of these disciplines were included in the study, if they were 18 years or older and already hospitalized in one of the study wards on the first day of the study. Excluded were patients with end-of-life care, communication difficulties such as not understanding or speaking the German language sufficiently to answer questions, with impaired communication due to aphasia or coma, or with severe hearing impairment.

Variables and measurements

Demographic information was collected on sex (male, female), age (in years), neurocognitive disorders (dementia, delirium), day of admission and the unit in which the patients were hospitalized. Delirium screening was exclusively performed with the mCAM-ED. The 4AT rating was derived from the items of the mCAM-ED. Corresponding items of the mCAM-ED were used to rate the 4AT algorithm (Table 1).

The 4 'A's Test (4AT)

The 4AT examines three cognitive domains (alertness, cognition, and attention), the factors incidence and course of symptoms:

- (a) Changes in alertness are assessed during the interview with the patient (MacLulich et al. 2011);

- (b) Changes in cognition of patients are assessed by the Abbreviated Mental Test-4 (AMT-4), a brief cognitive test with four questions on age, date of birth, current year and date (Schofield et al. 2010);
- (c) Changes in patient's attention are examined with the Month of the Year Backward Test (MBT) (Meagher et al. 2015), reciting months of the year from December to January (until July or further) (MacLulich et al. 2011); and
- (d) Acute changes in alertness, cognition and other cognitive functions in the previous two weeks or fluctuations in these areas in the previous 24 h are assessed using information from third parties (i.e. authorised representatives) and medical history (MacLulich et al. 2011).

Table 1 Variables of the mCAM-ED and the 4AT, listed according to the domains

Domains	mCAM-ED	4AT
Alertness	mRASS -5 (unarousable) to +4 (combative); alert & calm = 0	observation alert & calm (0) or clearly abnormal (4)
Cognition	MSQ	AMT-4
Orientation	10 questions (score: ≥ 3 errors indicate altered cognition)	4 questions (score: 0 errors = 0; 1 error = 1, ≥ 2 errors = 2)
To place	What is the name of this place? Where is this located (address)?	Place (name of hospital / building) –
To date	What is today's date? What is the month now? What is the year?	– – Current year
Towards oneself	How old are you? When were you born (month)? When were you born (year)?	Age Date of birth
On general information	Who is the president of the United States? Can you name a previous president?	– –
Disorganized thinking	Comprehension Test 4 questions (score: ≥ 2 mistakes indicate disorganized thinking) Will a stone float on water? Can you use a hammer to pound nails? Do two pounds of apples weigh more than one? Are there fish in the sea?	– – – –
Attention	MBT Omissions and time required for the task (score: ≥ 3 points indicate inattention)	MBT until July or further (score: ≥ 7 months = 0, starts but < 7 months / refuses = 1, untestable = 2)
Acute onset or fluctuation	Observation during the interview and information from third party	In the previous 2 weeks or 24 h (score: yes = 4; no = 0)

4AT the 4 'A's Test, mCAM-ED the modified Confusion Assessment Method for the Emergency Department, mRASS the Modified Richmond Agitation Sedation Scale, MSQ the Mental Status Questionnaire, AMT-4 the Abbreviated Mental Test 4, MBT the Months Backwards Test

According to the mCAM-ED delirium is considered when: (acute onset or fluctuation) and inattention and (disorganized thinking and/or alertness) are present. The 4AT total score is between 0 and 12. A score of ≥ 4 points indicates probable delirium

The score for item alertness and item acute onset or fluctuation is 0 or 4. The score for item AMT-4 and item attention is 0, 1 or 2. The cumulative score of the four sections is between 0 and 12. A value of 0 indicates that the probability of delirium in a patient is low. A score between 1 and 3 indicates a possible cognitive impairment of the tested patient. With a score of 4 or more, the presence of delirium is likely (MacLulich et al. 2011).

Since 2013, the 4AT has been validated in a variety of environments and patient populations: acute elderly emergency department (Gagne et al. 2018; Hendry et al. 2016; O’Sullivan et al. 2018), stroke units (Infante et al. 2017; Lees et al. 2013), geriatric wards (De et al. 2017; Bellelli et al. 2014; MacLulich et al. 2019), rehabilitation, orthogeriatric (De et al. 2017) and general medical wards (MacLulich et al. 2019; Kuladee and Prachason 2016). When validating against the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders IV Text Revision (DSM-IV-TR) or the Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5), the 4AT had a sensitivity of 76% to 93% and a specificity of 69.5–95% (De et al. 2017; Bellelli et al. 2014; Hendry et al. 2016; O’Sullivan et al. 2018; Infante et al. 2017; MacLulich et al. 2019; Kuladee and Prachason 2016). Sensitivity of the 4AT ranged from 84 to 100% and specificity from 74 to 82% compared to the CAM for delirium detection (Gagne et al. 2018; Lees et al. 2013). The assessment with the 4AT takes about two minutes on average (MacLulich et al. 2011).

The modified Confusion Assessment Method for the Emergency Department (mCAM-ED)

To assess delirium with mCAM-ED, changes in attention, cognition, disorganized thinking, and acute onset and/or fluctuation are assessed. The mCAM-ED consists of a two-step approach. This may be time saving in the ED setting. When used on wards for delirium consultation by the geriatric consultation team, however, both steps can be used on a regular basis (Hasemann 2019). In the first step, inattention is identified with the MBT (Meagher et al. 2015). In the second step, in cases of conspicuous attention, acute cognitive changes are evaluated by a structured interview with the Mental Status Questionnaire (MSQ) (Kahn et al. 1960), altered level of consciousness measured with the modified Richmond Agitation Sedation Scale (Chester et al. 2012) and disorganized thinking with the Comprehension Test (Hart et al. 1996). In addition, fluctuating symptoms are observed in the assessment. Possible or probable delirium is present when: (a) an acute change in cognition (MSQ) or observed fluctuating symptoms and (b) inattention (MBT) and (c) disorganized thinking and/or (d) a changed level of consciousness is identified in the interview (Fig. 1) (Inouye et al. 1990).

The mCAM-ED was developed and validated in the emergency department (Grossmann et al. 2014; Hasemann et al. 2018a). The mCAM-ED had a sensitivity of 90% (Confidence Interval [CI] 0.70; 0.97) and specificity of 98% (CI 0.95; 0.99) compared to DSM-IV-TR in delirium detection. The assessment with the mCAM-ED needed on average five minutes (Hasemann et al. 2018a).

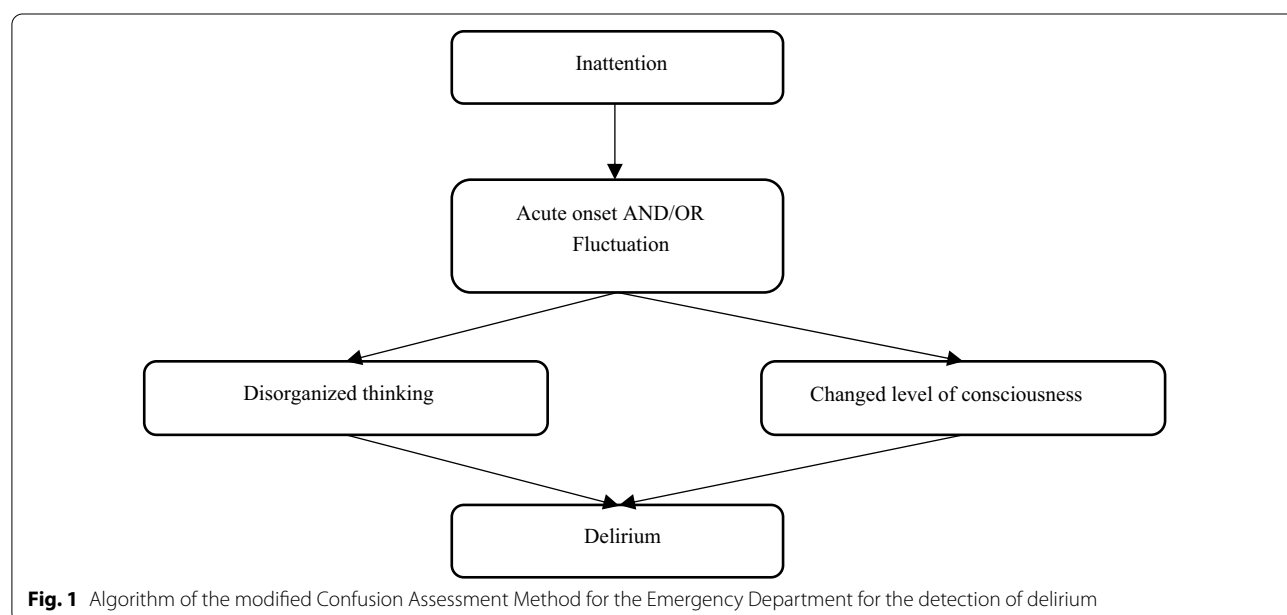


Fig. 1 Algorithm of the modified Confusion Assessment Method for the Emergency Department for the detection of delirium

Data collection and management

Data collection took place on two days in August 2018. On the first day of the study, the 32 trained nursing research assistants (RA) were assigned to eligible patients. The RAs visited the assigned patients on the first day of the study and conducted the assessments. On the second day of the study, the RAs supplemented data from the electronic medical record on the paper Case Report Form (CRF). This included demographic information such as sex, age, diagnosis of pre-existing dementia or recent delirium. The double entry method was done to minimize erroneous entries and increase data quality. In addition, any paper CRF without clear answers was discussed with the principal investigator to reach a consensus on the data. Using the corresponding items of the mCAM-ED, the 4AT algorithm was rated.

Training of research assistants

The training took place one day before the data collection. This training consisted of a one-day course on the screening and assessment of delirium and dementia. Case vignettes and video examples accompanied the theoretical input on delirium and dementia. An advanced practice nurse qualified as a PhD from the geriatric consultation service provided the training.

Ethical considerations

The RAs asked the patients to sign the written informed consent to participate in the study on the first day of data collection. If a patient was unable to make an informed decision, relatives or proxies were contacted to sign the written informed consent. The Ethics Committee Northwest Central Switzerland has approved this study under the number EKNZ: 2018-00,616.

Data analysis

Demographic information was analysed by frequencies and percentages for categorical variables. For continuous data, central tendencies with median and interquartiles (Q1, Q3) were reported. For the group comparisons, the Fisher's exact test or Mann-Whitney-U-test were performed according to the distribution of the data. These analyses were performed using IBM SPSS (Statistical Package for the Social Sciences) version 22.0. In order to address the aim, the performance of the 4AT, sensitivity, specificity, positive and negative likelihood ratio and the positive and negative predictive values (PPV and NPV) were calculated and presented together with a two-sided 95% CI. The exact CI for sensitivity, specificity, PPN and NPV was computed using Collett's method (Collett 1999). For the positive and negative likelihood ratio, a two-sided 95% CI was calculated and presented according to

Simel's method (Simel et al. 1991). The determination of the sample size was based on the validation study of the mCAM-ED (Hasemann et al. 2018a). Analysis of the performance was conducted using the software program "R" Version 3.6.1 (R Core Team 2018) with the package "epiR" (Stevenson 2020). Missing data were handled by pairwise deletion. The significance level for all analyses was set to $\alpha < 0.05$.

Results

Patient characteristics

During the study period, 211 patients were hospitalized on the six wards, of whom 116 were included in the study. A total of 95 patients were excluded due to, for instance, refusal to participate ($n=29$), admission after assessment ($n=22$) or absence of patients ($n=19$) (Fig. 2).

The median age of patients was 73.0 years (interquartile Q₁ 56.3, Q₃ 83.0) and 49 (42.2%) of the patients were women. Pre-existing dementia was present in 11 (9.5%) patients. A total of 8 out of 116 patients (6.9%) had delirium according to the gold standard, the mCAM-ED, and 16 (13.8%) had delirium according to the 4AT. Subgroup analysis of patients with delirium and patients without delirium showed no statistically significant difference in the development of delirium in respect to age ($U = -1.799$, $p = 0.072$), sex ($p = 1.000$) or pre-existing dementia ($p = 0.167$) (Table 2). There was a wide range of main diagnoses and heterogeneity among the patients (Additional file 1: Table 5), and two-thirds of them were hospitalized on medical wards (Table 2).

Performance

With the 4AT 100% (95% CI 0.63, 1.00) of all patients with delirium were correctly identified as being delirious when screened by RAs with the mCAM-ED (sensitivity). Additionally, the 4AT correctly identified 93% (95% CI 0.86, 0.97) of non-delirious patients as being

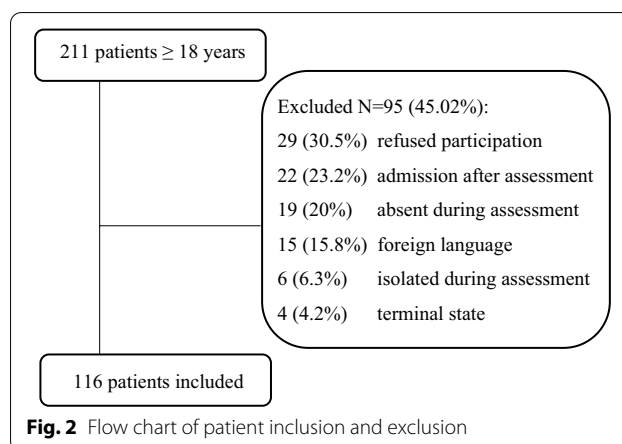


Table 2 Patient characteristics divided into delirium and no delirium

	Total sample		Delirium		No delirium		P value
	N = 116	%	N = 8	6.9%*	N = 108	93.1%	
<i>Age in years[‡]</i>							.072
Median (IQ)	73.0 (56.3; 83.0)		83.0 (62.3; 87.0)		72.5 (56.3; 82.8)		
Range	19–98		54–96		19–98		
<i>Sex[†]</i>							1.000
Female	49	42.2%	3	37.5%	46	42.6%	
Male	67	57.8%	5	62.5%	62	57.4%	
<i>Neurocognitive impairment[‡]</i>							.167
Dementia	11	9.5%	2	25.0%	9	8.3%	
No dementia	105	90.5%	6	75.0%	99	91.7%	
<i>Ward[‡]</i>							.345
Surgical wards	39	33.6%	1	12.5%	38	35.2%	
Medical wards	77	66.4%	7	87.5%	70	64.8%	
<i>Medications</i>							
Neuroleptics [†]	26	22.4%	3	37.5%	23	21.3%	.376
Benzodiazepines [†]	25	21.6%	1	12.5%	24	22.2%	1.000
Z-Drugs [†]	18	15.5%	4	50.0%	14	13.0%	0.020

IQ interquartile Q₁, Q₃, % percentage, [‡]Mann–Whitney-U-Test, [†]Fishers exact Test, *assessed with the modified Confusion Assessment Method for the Emergency Department (gold standard)

Table 3 Cross table of the 4AT results based on the mCAM-ED assessments

4AT	mCAM-ED		
	Delirium	No delirium	Total
Delirium	8 (6.9%)	8 (6.9%)	16 (13.8%)
No delirium	0 (0%)	100 (86.2%)	100 (86.2%)
Total	8 (6.9%)	108 (93.1%)	116 (100%)

4AT the 4 'A's Test, mCAM-ED the modified Confusion Assessment Method for the Emergency Department

non-delirious in comparison to the gold standard (specificity). Among delirium assignments of the 4AT, 50% (95% CI 0.25, 0.75) were judged as delirious by the RAs with the mCAM-ED (positive predictive value) and 100% (95% CI 0.96, 1.00) of the 4AT non-delirium assignments were confirmed by the gold standard (negative predictive value). The probability of a patient in delirium being correctly classified as delirious with the 4AT was 13.5 times (95% CI 6.93, 26.30) higher than in a non-delirious patient falsely classified as delirious (positive likelihood ratio). Vice versa, the probability of falsely classifying a patient in delirium as non-delirious with the 4AT was zero (negative likelihood ratio) (Tables 3, 4).

For both tools, there were variations in the impact of the scores. In two cases the assessment of attention was considered to be an attention deficit, in three cases not. This was the consequence of different cut-offs rating

Table 4 Performance of the 4AT in delirium detection compared to the mCAM-ED

	Total sample n = 116	
	Point estimates	95% CI
Sensitivity	1.00	(0.63, 1.00)
Specificity	0.93	(0.86, 0.97)
Positive predictive value	0.50	(0.25, 0.75)
Negative predictive value	1.00	(0.96, 1.00)
Positive likelihood ratio	13.50	(6.93, 26.30)
Negative likelihood ratio	0.00	(0.00, NaN)

CI confidence interval, 4AT the 4 'A's Test, mCAM-ED the modified Confusion Assessment Method for the Emergency Department; NaN not a number

inattention by the 4AT and the mCAM-ED (Additional file 2: Table 6).

Discussion

In this comparison of the 4AT with the mCAM-ED in detecting delirium in hospitalized patients, the 4AT showed good performance, but our expectations of a 95% agreement were not fulfilled: The 4AT and the mCAM-ED did not perform equally well. The false-positive rate of the 4AT was exactly twice. The 4AT requires only an impaired level of alertness or an acute onset or fluctuation in order to suggest the presence of delirium, whereas the algorithm of the mCAM-ED requires more items to positively rate a delirium.

Although the mCAM-ED consists of more items and an additional neurocognitive domain (disorganized thinking), the remaining neurocognitive domains (attention, cognition and alertness) do not differ between the mCAM-ED and the 4AT. However, the weighting of the domains in the algorithm for detecting delirium varies between the mCAM-ED and the 4AT. While three out of four domains of the mCAM-ED are required to confirm delirium (Fig. 1) (Grossmann et al. 2014; Hasemann et al. 2018a), the 4AT requires only one domain (i.e. acute onset and/or fluctuation or alertness) (MacLulich et al. 2011). In the end, the 4AT increased its sensitivity at the expense of specificity by using a single criterion to confirm delirium. It is noteworthy that the mCAM-ED is closer to the DSM criteria than the 4AT, which is due to the use of three items instead of one criterion, as in the case of the 4AT. According to DSM-5, the following five points for the delirium diagnosis must be fulfilled: (a) disturbed attention or awareness, (b) sudden change and fluctuation in severity, (c) an additional disturbance in cognition, (d) not better explained by the presence of a neurocognitive disorder, and (e) context of medical history, substance abuse or withdrawal, or exposure to a toxin (American Psychiatric Association 2013). To meet the DSM-5 criteria of delirium, the extent of coverage of the mCAM-ED criteria is greater than that of the 4AT. Thus, the 4AT is more susceptible to overestimating the presence of delirium than the mCAM-ED.

Considering the different sample and settings, in our study the 4AT has a similarly good performance as in the first validation study by Bellelli et al. (2014). In their study in comparison to DSM-IV-TR, sensitivity and specificity of 89.7% and 84.1%, respectively (Bellelli et al. 2014), show a slight difference to this comparison with the mCAM-ED, where sensitivity and specificity were 100% and 93%, respectively. In the recently published validation study in patients in the emergency department or acute general wards, the 4AT showed a lower sensitivity of 75% and a higher specificity of 94.5% compared to the DSM-IV-TR (MacLulich et al. 2019). In a validation study of several short screening tools compared to the CAM in a stroke unit, the 4AT showed a good sensitivity of 100% and a reasonable specificity of 82% (Lees et al. 2013).

Delirium prevalence of 6.9% measured in this study with the mCAM-ED is consistent with the data from other Swiss studies. Similar prevalence of 7.0 and 9.5%, respectively, was measured with DSM-IV-TR (Hasemann et al. 2018a) and the mCAM-ED in the emergency department (Grossmann et al. 2014). In a pilot study conducted in a central hospital, the prevalence was twice as high (14%) as in this study (Schwarber et al. 2017). However, in a recently published study, the prevalence

of delirium in intermediate/general medicine wards was nearly four times higher (27.3%) than in this study (Schubert et al. 2018). These figures demonstrate that the prevalence of delirium in Switzerland is highly variable. One aspect contributing to variability in prevalence is the method of measurement. In the hospital-wide Swiss study by Schubert et al. (2018), a period prevalence was measured with the Delirium Observation Scale (DOS) (Schubert et al. 2018). However, given some shortcomings of the DOS to discriminate between delirium and dementia (specificity 92%) (Hasemann et al. 2018b), this might explain the high prevalence. The unknown proportion of patients with dementia in the study by Schubert et al. (2018) might possibly have contributed to a high prevalence of delirium by misclassifying dementia as delirium. Furthermore, the period measurement could also be a contribution. In contrast to other studies, Schubert et al. (2018) determined a period prevalence and not a point prevalence. Since delirium does not usually occur over the entire hospitalization period, but only over several days, a period prevalence would capture more delirium than a point prevalence. In another Swiss study, in which a prevalence of around 16% was measured throughout the entire hospital stay in the medical wards, a considerable variability in prevalence was observed from day to day. Delirium point prevalence varied from 15% to about 50% between days in comparison with the prevalence across the entire hospital stay in same patient sample (Hasemann 2014). Symptoms of delirium are not observed constantly, but are subject to fluctuations over the course of the day (American Psychiatric Association 2013). It is clear that a point prevalence only reflects the condition at one specific time period and is therefore subject to fluctuations. These need be taken into account in the analysis, and so our measured prevalence is subject to these fluctuations.

As in the first and in a recently published validation study of the 4AT, patients aged 70 years and older were included (Bellelli et al. 2014; MacLulich et al. 2019). On average, the patients of Bellelli et al. (2014) were 84 years old (Bellelli et al. 2014), which is considerably higher than in our sample. In a point prevalence study conducted in Ireland, almost 20% of all inpatients with a median age of 69 years had delirium (Ryan et al. 2013). This is considerably higher than in the sample in this study despite the similar median age. The variability in the inclusion criterion of age on the one hand and the contextual setting and population on the other hand could explain the differences in prevalence. Patients aged 65 and older are more predisposed to the development of delirium due to their age (Inouye 2006) and in intensive care units a markedly higher prevalence and incidence of delirium is observed than in regular wards (Inouye et al. 2014).

However, depending on the specialisation of regular wards, different incidences of delirium may occur in relation to patient populations, with surgical wards having a higher incidence than medical wards (Inouye et al. 2014).

Compared to the gold standard, the 4AT classified twice as many patients as delirious. The false-positive cases may generate more in-depth assessments, resulting in an increased burden for patients and nursing staff. Nurses are already challenged by the increased demand for care and the increased workload of delirious patients (Schubert et al. 2018), so that further in-depth assessments of false-positive results might lead to an increase in the already high workload. This may result in a lower acceptance of delirium screening with high false-positive rates.

Several studies claim that the 4AT is a brief and easy to use tool that requires no training (De and Wand 2015; MacLulich et al. 2011; Bellelli et al. 2014). Bellelli et al. (2014) pointed out that experienced physicians used the 4AT and that other professionals such as nurses should test this tool in everyday clinical practice (Bellelli et al. 2014). However, in a quality improvement project on an acute geriatric ward, the 4AT showed a low performance (sensitivity 50.0%, specificity 86.2%). In that project, the nurses were given a lecture with information about the tool and a two-week training course (Myrstad et al. 2019). It was noted that the sensitivity of the application of the 4AT in daily clinical routine was low, despite short practicing phase. Therefore, in-depth training for staff prior to the introduction of the 4AT in the daily clinical routine was suggested (Myrstad et al. 2019). Additionally, similar findings were reported in another quality improvement project in a hospice, where the 4AT was applied to patients at admission. In spite of initial implementation challenges, the tool was seen as useful in this population. However, the study concluded that an increased use of the screening requires continuous feedback and training (Baird and Spiller 2017). It illustrates the need for awareness and training for early and appropriate detection. Main issues of the low performance were nurses' difficulties in rating acute onset and/or fluctuation (Myrstad et al. 2019). This illustrates, even with easy-to-use tools, the need to train for adequate application, but also for the detection of delirium, so that early intervention can be carried out and negative consequences avoided.

Strengths and limitations

A strength of this study is the fact that the ratings of the mCAM-ED and the 4AT originate from the same person, therefore no inter-rater differences are present. Another strength of this study is the consecutive sampling to minimize potential sampling bias. In accordance

with the approach of Voyer et al. (2008) to use the CAM as a gold standard in the accuracy study for delirium symptoms based on the entries in the nurses' medical records (Voyer et al. 2008), our approach to use the mCAM-ED as a gold standard corresponds to current practice and is valid. The CAM was also used as a gold standard in two 4AT validation studies (Gagne et al. 2018; Lees et al. 2013). Certain shortcomings of this study must be acknowledged. Only 116 of 211 inpatients could be included in this pilot study. The high rejection rate may have been contributed to by less experienced and young RAs. The RAs may have felt overwhelmed and did not dare to contact patients' relatives to seek informed consent, so that potentially vulnerable patients were not included. In addition, many young people do not like talking on the phone (Kupke 2020), which may have been another inhibition threshold. Another limitation is the low prevalence of delirium at 6.9%. The low prevalence may result from a consecutive sample with a low median age of 73 years and the assessment of delirium on a point prevalence basis. Furthermore, the low prevalence of delirium may be due to the nurse-led hospital-integrated delirium consultation service and an established delirium management programme. Moreover, our sample is very heterogeneous with an age range from 19 to 98 years and also includes patients with a diagnosed neurocognitive disorder. Initially, we planned to conduct a sub-analysis with stratified age groups. However, the sample was too low to do so.

Conclusions

In this first comparison in delirium detection in inpatients, the 4AT showed a high false-positive rate of delirium assessments. This may result in an increased need for further in-depth assessments, which may contribute to higher workload and higher screening burden of the patient. Although the 4AT claims to require no training, nurses need more support, as our study and other studies suggest.

Abbreviations

4AT: The 4 'A's test; AMT-4: The abbreviated mental test-4; CAM: The confusion assessment method; CI: Confidence interval; CRF: Case report form; DOS: Delirium observation scale; DSM: Diagnostic and statistical manual of mental disorders; DSM-IV-TR: Diagnostic and statistical manual of mental disorders IV text revision; DSM-5: Diagnostic and statistical manual of mental disorders fifth edition; IQ: Interquartile; MBT: The months of the year backwards test; mCAM-ED: The modified confusion assessment method for the emergency department; mRASS: The modified Richmond agitation sedation scale; MSQ: The mental status questionnaire; RA: Research assistant; SD: Standard deviation; PPV: Positive predictive value; NPV: Negative predictive value; +LR: Positive likelihood ratio; -LR: Negative likelihood ratio.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42269-022-00863-4>.

Additional file 1: Table 5. Characteristics of the patients: Main diagnoses according to ICD-10 divided into delirium and no delirium.

Additional file 2: Table 6. Differences in delirium screening: List of algorithms and false-positive ratings of 4AT per participant.

Acknowledgements

The authors thank Manuela Pretto and Anja Hermann for their support in collecting data. Furthermore, the authors thank Sabrina Klimmeck, Luisa Eggenschwiler, Lili Saar, Simone Stamm and Pia Urfer for their stimulating discussions. Additional thanks go to Nancy Hänsler for the linguistic proofreading of the text.

Author contributions

WH had the idea for the study, WH, KS, OM, FZ, SB and BG drafted the study design, WH and KS performed the data collection, KS developed the data matrix, OM carried out the data management and input, WH and OM performed the statistical analysis and interpretation of the results, OM prepared the manuscript, SB, BG, WH, FZ critically revised the manuscript. All authors have read and approved the manuscript.

Funding

There were no specific grants or funding for this research.

Availability of data and materials

Not applicable.

Declarations

Ethical approval and consent to participate

The Ethics Committee Northwest Central Switzerland has approved the study (EKNZ: 2018-00616). The research assistants asked the patients to sign the written informed consent to participate in the study on the first day of data collection. If a patient was not able to make an informed decision, relatives or proxy were contacted to sign the written informed consent.

Consent for publication

Not applicable.

Competing interests

The authors state that there is no conflict of interest.

Author details

¹Department Public Health, Faculty of Medicine, Institute of Nursing Science, University of Basel, Bernoullistrasse 28, 4056 Basel, Switzerland. ²Department Chest, Abdomen and Pelvis, University Hospital Basel, Spitalstrasse 21, 4031 Basel, Switzerland. ³Practice Development Unit, Department of Nursing and Allied Health Professions, University Hospital Basel, Hebelstrasse 2, 4031 Basel, Switzerland. ⁴Department of Nursing Science, University of Applied Sciences (HS-Gesundheit Bochum), Gesundheitscampus 6-8, 44801 Bochum, Germany. ⁵Development and Research Unit, University Department of Psychiatry, Wilhelm Klein-Strasse 27, 4002 Basel, Switzerland. ⁶Nursing Development and Research, University Department of Geriatric Medicine FELIX PLATTER, Burgfelderstrasse 101, 4002 Basel, Switzerland.

Received: 18 September 2021 Accepted: 11 June 2022

Published online: 21 June 2022

References

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (DSM-5®), vol 5. American Psychiatric Association, Arlington, VA

- Baird L, Spiller JA (2017) A quality improvement approach to cognitive assessment on hospice admission: could we use the 4AT or Short CAM? *BMJ Open Qual* 6(2):e000153. <https://doi.org/10.1136/bmjopen-2017-000153>
- Universitätsspital Basel (2019) Geschäftsbericht 2018. Universitätsspital Basel, Basel
- Belanger L, Ducharme F (2011) Patients' and nurses' experiences of delirium: a review of qualitative studies. *Nurs Crit Care* 16(6):303–315. <https://doi.org/10.1111/j.1478-5153.2011.00454.x>
- Bellelli G, Morandi A, Davis DH, Mazzola P, Turco R, Gentile S, Ryan T, Cash H, Guerini F, Torpilliesi T, Del Santo F, Trabucchi M, Annoni G, MacLulich AM (2014) Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing* 43(4):496–502. <https://doi.org/10.1093/ageing/afu021>
- Chester JG, Beth Harrington M, Rudolph JL, Group VADW (2012) Serial administration of a modified Richmond Agitation and Sedation Scale for delirium screening. *Journal of hospital medicine* 7(5):450–453. <https://doi.org/10.1002/jhm.1003>
- Cohen MZ, Pace EA, Kaur G, Bruera E (2009) Delirium in advanced cancer leading to distress in patients and family caregivers. *J Palliat Care* 25(3):164–171. <https://doi.org/10.1177/082585970902500303>
- Collett D (1999) Modelling binary data. Chapman & Hall/CRC, Boca Raton, FL
- Collins N, Blanchard MR, Tookman A, Sampson EL (2010) Detection of delirium in the acute hospital. *Age Ageing* 39(1):131–135. <https://doi.org/10.1093/ageing/afp201>
- Davis DH, Muniz-Terrera G, Keage HA, Stephan BC, Fleming J, Ince PG, Matthews FE, Cunningham C, Ely EW, MacLulich AM, Brayne C, Epidemiological Clinicopathological Studies in Europe Collaborative M (2017) Association of delirium with cognitive decline in late life: a neuropathologic study of 3 population-based cohort studies. *JAMA psychiatry* 74(3):244–251. <https://doi.org/10.1001/jamapsychiatry.2016.3423>
- De J, Wand AP (2015) Delirium screening: a systematic review of delirium screening tools in hospitalized patients. *Gerontologist* 55(6):1079–1099. <https://doi.org/10.1093/geront/gnv100>
- De J, Wand APF, Smerdely PI, Hunt GE (2017) Validating the 4A's test in screening for delirium in a culturally diverse geriatric inpatient population. *Int J Geriatr Psychiatry* 32(12):1322–1329. <https://doi.org/10.1002/gps.4615>
- Fick DM, Inouye SK, Guess J, Ngo LH, Jones RN, Saczynski JS, Marcantonio ER (2015) Preliminary development of an ultrabrief two-item bedside test for delirium. *J Hosp Med* 10(10):645–650. <https://doi.org/10.1002/jhm.2418>
- Gagne AJ, Voyer P, Boucher V, Nadeau A, Carmichael PH, Pelletier M, Gouin E, Berthelot S, Daoust R, Wilchesky M, Richard H, Pelletier I, Ballard S, Lague A, Emond M (2018) Performance of the French version of the 4AT for screening the elderly for delirium in the emergency department. *CJEM* 20(6):903–910. <https://doi.org/10.1017/cem.2018.367>
- Grossmann FF, Hasemann W, Graber A, Bingisser R, Kressig RW, Nickel CH (2014) Screening, detection and management of delirium in the emergency department - a pilot study on the feasibility of a new algorithm for use in older emergency department patients: the modified Confusion Assessment Method for the Emergency Department (mCAM-ED). *Scand J Trauma Resusc Emerg Med* 22(1):19. <https://doi.org/10.1186/1757-7241-22-19>
- Hasemann W (2014) Effects of the interdisciplinary basel dementia and delirium management program DemDel. In: Paper presented at the the 24th Alzheimer Europe Conference, Glasgow, Scotland, UK 20–22 October 2014
- Hasemann W (2019) Delirium Consulting Service at the University Hospital Basel
- Hasemann W, Grossmann FF, Stadler R, Bingisser R, Breil D, Hafner M, Kressig RW, Nickel CH (2018a) Screening and detection of delirium in older ED patients: performance of the modified Confusion Assessment Method for the Emergency Department (mCAM-ED). A two-step tool. *Intern Emerg Med* 13 (1970–9366 (Electronic)):915–922. <https://doi.org/10.1007/s11739-017-1781-y>
- Han JH, Vasilevskis EE, Chandrasekhar R, Liu X, Schnelle JF, Dittus RS, Ely EW (2017) Delirium in the emergency department and its extension into hospitalization (DELINATE) study: effect on 6-month function and cognition. *J Am Geriatr Soc* 65(6):1333–1338. <https://doi.org/10.1111/jgs.14824>
- Hart RP, Levenson JL, Sessler CN, Best AM, Schwartz SM, Rutherford LE (1996) Validation of a cognitive test for delirium in medical ICU patients. *Psychosomatics* 37(6):533–546. [https://doi.org/10.1016/S0033-3182\(96\)71517-7](https://doi.org/10.1016/S0033-3182(96)71517-7)

- Hasemann W, Tolson D, Godwin J, Spirig R, Frei IA, Kressig RW (2018b) Nurses' recognition of hospitalized older patients with delirium and cognitive impairment using the delirium observation Screening scale: a prospective comparison study. *J Gerontol Nurs* 44(12):35–43. <https://doi.org/10.3928/00989134-20181018-02>
- Hendry K, Quinn TJ, Evans J, Scortichini V, Miller H, Burns J, Cunnington A, Stott DJ (2016) Evaluation of delirium screening tools in geriatric medical inpatients: a diagnostic test accuracy study. *Age Ageing* 45(6):832–837. <https://doi.org/10.1093/ageing/afw130>
- Infante MT, Pardini M, Balestrino M, Finocchi C, Malfatto L, Bellelli G, Mancardi GL, Gandolfo C, Serrati C (2017) Delirium in the acute phase after stroke: comparison between methods of detection. *Neurol Sci: off J Ital Neurol Soc Ital Soc Clin Neurophysiol* 38(6):1101–1104. <https://doi.org/10.1007/s10072-017-2832-x>
- Inouye SK (2006) Delirium in older persons. *N Engl J Med* 354(11):1157–1165. <https://doi.org/10.1056/NEJMra052321>
- Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI (1990) Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med* 113(12):941–948. <https://doi.org/10.7326/0003-4819-113-12-941>
- Inouye SK, Foreman MD, Mion LC, Katz KH, Cooney LM Jr (2001) Nurses' recognition of delirium and its symptoms: comparison of nurse and researcher ratings. *Arch Intern Med* 161(20):2467–2473. <https://doi.org/10.1001/archinte.161.20.2467>
- Inouye SK, Westendorp RG, Saczynski JS (2014) Delirium in elderly people. *Lancet* 383(9920):911–922. [https://doi.org/10.1016/S0140-6736\(13\)60688-1](https://doi.org/10.1016/S0140-6736(13)60688-1)
- Kahn RL, Goldfarb AI, Pollack M, Peck A (1960) Brief objective measures for the determination of mental status in the aged. *Am J Psychiatry* 117(4):326–328. <https://doi.org/10.1176/ajp.117.4.326>
- Kupke S (2020) Die Angst vor dem Telefonieren wächst. *Forschung & Lehre Kuladee S, Prachason T (2016) Development and validation of the Thai version of the 4 'A's Test for delirium screening in hospitalized elderly patients with acute medical illnesses. Neuropsychiatr Dis Treat* 12:437–443. <https://doi.org/10.2147/NDT.S97228>
- Lees R, Corbet S, Johnston C, Moffitt E, Shaw G, Quinn TJ (2013) Test accuracy of short screening tests for diagnosis of delirium or cognitive impairment in an acute stroke unit setting. *Stroke* 44(11):3078–3083. <https://doi.org/10.1161/strokeaha.113.001724>
- Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK (2008) One-year health care costs associated with delirium in the elderly population. *Arch Intern Med* 168(1):27–32. <https://doi.org/10.1001/archinternmed.2007.4>
- MacLulich A, Ryan A, Cash H (2011) The 4 'A's Test <https://www.the4at.com/>
- MacLulich AM, Shenkin SD, Goodacre S, Godfrey M, Hanley J, Stiobhairt A, Lavender E, Boyd J, Stephen J, Weir C, MacRaild A, Steven J, Black P, Diernberger K, Hall P, Tiegies Z, Fox C, Anand A, Young J, Siddiqi N, Gray A (2019) The 4 'A's test for detecting delirium in acute medical patients: a diagnostic accuracy study. *Health Technol Assess* 23(40):1–194. <https://doi.org/10.3310/hta23400>
- Meagher J, Leonard M, Donoghue L, O'Regan N, Timmons S, Exton C, Cullen W, Dunne C, Adams D, MacLulich AJ, Meagher D (2015) Months backward test: a review of its use in clinical studies. *World J Psychiatry* 5(3):305–314. <https://doi.org/10.5498/wjp.v5.i3.305>
- Mittal V, Muralee S, Williamson D, McEnerney N, Thomas J, Cash M, Tampi RR (2011) Review: delirium in the elderly: a comprehensive review. *Am J Alzheimers Dis Other Demen* 26(2):97–109. <https://doi.org/10.1177/1533317510397331>
- Moskowitz EE, Overbey DM, Jones TS, Jones EL, Arcomano TR, Moore JT, Robinson TN (2017) Post-operative delirium is associated with increased 5-year mortality. *Am J Surg* 214(6):1036–1038. <https://doi.org/10.1016/j.amjsurg.2017.08.034>
- Myrstad M, Watne LO, Johnsen NT, Bors-Lind E, Neerland BE (2019) Delirium screening in an acute geriatric ward by nurses using 4AT: results from a quality improvement project. *Eur Geriatr Med* 10(4):667–671. <https://doi.org/10.1007/s41999-019-00215-y>
- Olofsson B, Persson M, Bellelli G, Morandi A, Gustafson Y, Stenvall M (2018) Development of dementia in patients with femoral neck fracture who experience postoperative delirium-A three-year follow-up study. *Int J Geriatr Psychiatry* 33(4):623–632. <https://doi.org/10.1002/gps.4832>
- O'Sullivan D, Brady N, Manning E, O'Shea E, O'Grady S, O'Regan N, Timmons S (2018) Validation of the 6-Item Cognitive Impairment Test and the 4AT test for combined delirium and dementia screening in older Emergency Department attendees. *Age Ageing* 47(1):61–68. <https://doi.org/10.1093/ageing/afx149>
- Ryan DJ, O'Regan NA, Caoimh RO, Clare J, O'Connor M, Leonard M, McFarland J, Tighe S, O'Sullivan K, Trzepacz PT, Meagher D, Timmons S (2013) Delirium in an adult acute hospital population: predictors, prevalence and detection. *BMJ Open*. <https://doi.org/10.1136/bmjopen-2012-001772>
- R Core Team (2018) R: a language and environment for statistical computing, 3.5.1 edn. R Foundation for Statistical Computing, Vienna, Austria
- Schofield I, Stott DJ, Tolson D, McFadyen A, Monaghan J, Nelson D (2010) Screening for cognitive impairment in older people attending accident and emergency using the 4-item Abbreviated Mental Test. *Eur J Emerg Med* 17(6):340–342. <https://doi.org/10.1097/MEJ.0b013e32833777ab>
- Stevenson M (2020) epiR: tools for the analysis of epidemiological data. 1.0–11 edn.
- Schubert M, Schurch R, Boettger S, Garcia Nunez D, Schwarz U, Bettex D, Jenewein J, Bogdanovic J, Staehli ML, Spirig R, Rudiger A (2018) A hospital-wide evaluation of delirium prevalence and outcomes in acute care patients - a cohort study. *BMC Health Serv Res* 18(1):550. <https://doi.org/10.1186/s12913-018-3345-x>
- Schwarber A, Hasemann W, Stillhard U, Schoop B, Senn B (2017) Prevalence, influencing factors, screening and diagnosis of delirium at stationary hospitalised patients in a Swiss acute Hospital. *A Pilot Study Pflege* 30(4):199–208. <https://doi.org/10.1024/1012-5302/a000564>
- Simel DL, Samsa GP, Matchar DB (1991) Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 44(8):763–770. [https://doi.org/10.1016/0895-4356\(91\)90128-V](https://doi.org/10.1016/0895-4356(91)90128-V)
- Voyer P, Cole MG, McCusker J, St-Jacques S, Laplante J (2008) Accuracy of nurse documentation of delirium symptoms in medical charts. *Int J Nurs Pract* 14(2):165–177. <https://doi.org/10.1111/j.1440-172X.2008.00681.x>
- Wand AP, Thoo W, Ting V, Baker J, Sciuriaga H, Hunt GE (2013) Identification and rates of delirium in elderly medical inpatients from diverse language groups. *Geriatr Nurs* 34(5):355–360. <https://doi.org/10.1016/j.gerinurse.2013.05.004>
- Wei LA, Fearing MA, Sternberg EJ, Inouye SK (2008) The confusion assessment method: a systematic review of current usage. *J Am Geriatr Soc* 56(5):823–830. <https://doi.org/10.1111/j.1532-5415.2008.01674.x>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)