The observational scale of level of arousal: a brief tool for assessing and monitoring level of arousal in patients with delirium outside the ICU

Hall, Roanna; Stiobhairt, Antaine; Allerhand, Mike; MacLullich, Alasdair M.J.; Tieges, Zoe

Published in:
International Journal of Geriatric Psychiatry

DOI:
10.1002/gps.5324

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Received: 12 October 2019 | Accepted: 25 April 2020
DOI: 10.1002/gps.5324

RESEARCH ARTICLE

International Journal of Geriatric Psychiatry WILEY

The Observational Scale of Level of Arousal: A brief tool for assessing and monitoring level of arousal in patients with delirium outside the ICU

Roanna Hall[1,2,3] | Antaine Stíobhairt[1] | Mike Allerhand[2] | Alasdair M. J. MacLullich[1,2,3] | Zoë Tieges[1,2]

[1]Edinburgh Delirium Research Group, University of Edinburgh, Edinburgh, UK

[2]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

[3]Medicine of the Elderly Department, Royal Infirmary of Edinburgh, Edinburgh, UK

Correspondence
Zoë Tieges, Edinburgh Delirium Research Group, University of Edinburgh, Room S1642, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK.
Email: zoe.tieges@ed.ac.uk

Funding information
Research into Ageing (AgeUK) and the British Geriatrics Society, Grant/Award Number: 342

Objectives: Altered level of arousal, encompassing drowsiness and hypervigilance, affects at least 10% of acutely unwell patients. Existing scales provide limited coverage of milder changes in level of arousal. We devised the Observational Scale of Level of Arousal (OSLA) to enable more detailed arousal assessment. Here, we provide a preliminary case-control study of performance of the OSLA in assessing abnormal level of arousal associated with delirium outside the ICU.

Methods: Hip fracture patients (N = 108, median age = 82 years) were assessed for delirium pre- and post-operatively using the Confusion Assessment Method and the Delirium Rating Scale-Revised-98. The OSLA has four graded items assessing eye opening, eye contact, posture, and movement (score range 0 [normal arousal]-15). We assessed the psychometric and diagnostic characteristics of the OSLA. Adjusted linear mixed effects models were used to explore responsiveness of the OSLA to within-patient change in delirium status.

Results: A total of 44 patients (40.7%) were diagnosed with delirium. OSLA scores were higher in delirium (pooled median = 3, InterQuartile Range [IQR] = 2-5) compared to no delirium (pooled median = 1, IQR = 1-2; *P*-values <.05 to <.001). The Area under the Receiver Operating Characteristic curve was 0.82 (95% Confidence Interval (CI) = 0.77-0.86). OSLA scores were responsive to change in delirium status (ß = −3.09. SE = 1.41, *P* < .03).

Conclusions: This study provides preliminary evidence supporting use of the OSLA as an instrument for identifying abnormal level of arousal associated with delirium and monitoring this longitudinally. Further validation in larger cohorts with blinded raters is required.

KEYWORDS
arousal, attention, cognition, delirium, orthopaedic surgery

Roanna Hall and Antaine Stíobhairt contributed equally to this study. Alasdair M. J. MacLullich and Zoë Tieges contributed equally to this study.

1 | INTRODUCTION

Delirium is a severe, acute neurocognitive disorder characterised by disturbances in attention, level of arousal and other mental functions. It affects at least one in eight hospitalised older patients and is independently associated with multiple adverse outcomes.[1-3]

Alterations in level of arousal are common in delirium, with many patients showing hypo- or hyperarousal.[4] The hypoactive subtype, characterised by drowsiness or somnolence, is the most common form of delirium.[5]

The arousal component of delirium has been described variably in standard diagnostic criteria (Table 1). In DSM-5, severely reduced level of arousal precluding cognitive testing or interview but above the level of coma is considered to indicate severe inattention. Assessment of level of arousal is therefore a core part of the evaluation of the features of delirium, and arousal measurements appear to be useful in clinical practice as a strong indicator of delirium.[4,6,7]

Outside the field of delirium, abnormal level of arousal is increasingly seen as a crucial marker of illness severity and predictor of mortality in hospitalised patients.[8] In United Kingdom hospitals, level of arousal is routinely assessed using the AVPU scale (A, alert; V, responds to voice; P, responds to pain; U, unresponsive) as one of six indicators as a National Early Warning Score.[9] Yet level of arousal in the specific context of delirium remains relatively understudied compared to its key cognitive symptom of attention deficits and other features.[10-13]

TABLE 1 Diagnostic and Statistical Manual of Mental Disorders (DSM) and International Classification of Diseases (ICD) diagnostic criteria descriptions of arousal disturbance in delirium

DSM/ICD	Description
DSM-III	Clouding of consciousness (reduced clarity of awareness of the environment), with reduced capacity to shift, focus, and sustain attention to environmental stimuli (Criterion A) Disturbance of sleep-wakefulness cycle, with insomnia or daytime drowsiness (one of four features under Criterion B, of which at least two must be present)
DSM-III-Revised	Reduced level of consciousness, for example, difficulty keeping awake during examination (one of six features under Criterion C, of which at least two must be present)
DSM-IV	A disturbance of consciousness, that is, reduced clarity of awareness of the environment, with reduced ability to focus, sustain, or shift attention (Criterion A)
DSM-V	A disturbance of attention (that is, reduced ability to direct, focus, sustain, and shift attention) and awareness (reduced orientation to the environment)
ICD-10	Clouding of consciousness, that is, reduced clarity of awareness of the environment, with reduced ability to focus, sustain, or shift attention (Criterion A)

Key points

- Assessment of level of arousal is a core part of the evaluation of delirium, and arousal measurements could be useful in clinical practice as a strong indicator of delirium.
- The Observational Scale of Level of Arousal (OSLA) was developed as a brief observational instrument to characterise the abnormalities of level of arousal associated with delirium (score range 0 [normal arousal]-15). It comprises four items evaluating different aspects of arousal: eye opening, eye contact, posture, and movement.
- The area under the Receiver Operating Characteristic curve for the OSLA for detecting delirium was 0.82. OSLA scores were responsive to within-patient change in delirium status and severity over time.
- This study provides support for the utility of the OSLA as a brief, accurate instrument for measuring level of arousal in delirium and for monitoring change in arousal in non-ICU patients over time. Further validation studies are necessary to establish the clinical utility of the OSLA.

In delirium, level of arousal is often reduced but there is a wide range of severities, from mild drowsiness to only being able to produce a basic motor response to a verbal stimulus. Conversely, patients may have heightened arousal and appear agitated and hyperalert. The Richmond Agitation-Sedation Scale (RASS),[14] which was originally developed to assess agitation or sedation levels in Intensive Care Unit (ICU) patients, has recently been modified for use as a delirium screen by including assessment of attention (mRASS).[7] The RASS is the most studied arousal scale in delirium.[4,15] However, a RASS score of +1 or −1 does not provide detailed information on the degree to which level of arousal is abnormal. More generally, an overall lack of granularity and operationalisation to capture this important feature of delirium in both arousal-specific and general delirium scales suggests that there would be value in having an instrument that provides a more detailed assessment.

To address this, we developed a new scale entitled the Observational Scale of Level of Arousal (OSLA; Table 2). It was designed for research use, to characterise the abnormalities of level of arousal associated with delirium, complementing assessments of attention and other features of delirium. Individual item scores characterise the profile of abnormalities while the single overall score provides an index of severity. The OSLA was used in a small study examining the relationship between altered arousal and inattention.[4] OSLA scores were strongly associated with delirium diagnosis. Another study reported good diagnostic accuracy of the

TABLE 2 The Observational Scale of Level of Arousal

Observational Scale of Level of Arousal (OSLA)
Eye opening
Score Description
0 Open on arrival and remain so, under patient's control, outlasts stimulus
1 Open on arrival but close if stimulus removed
1 Open to voice but then outlasts stimulus
2 Open to voice but close if stimulus removed
3 Open to gentle physical stimulation (squeezing hand, gently shaking shoulder)
4 Open to pain only
5 No eye opening
Eye contact
Score Description
0 Spontaneously makes and holds eye contact appropriately
1 Drowsy and makes eye contact to command but cannot hold it for very long
1 Alert but eyes wandering, some appropriate eye contact
2 Alert but eyes wandering, little or no appropriate eye contact
2 Drowsy but makes brief eye contact
3 Eyes will/are open but no eye contact
Posture (NB take into account weakness due to stroke or neurological disease, etc.)
Score Description
0 Sitting out in chair or up in bed, holding appropriate posture
1 Slumped in chair or bed but attempts to sit upright and sustain posture on request
2 Slumped in chair or bed and unable to sustain posture
3 Lying in bed and unable or no response to request to sustain posture
Movement
Score Description
0 Moves spontaneously and purposefully with no restless or agitated movements
1 Occasional or mild restless or fidgety movements, no aggressive or vigorous movements
1 Reduced frequency of movement, mildly slowed up
2 Frequent restless or fidgety movements, no aggressive or vigorous movements
2 Moderately reduced frequency and speed of movement, interfering with assessment or self-care
3 Aggressive or vigorous, recent pulling out of lines
4 Overtly combative, violent
4 Severely reduced frequency and speed of movement, few spontaneous movements
Score (0-15)

OSLA for detecting delirium on its own, and in combination with an attention task.[16]

Here, we provide preliminary evaluation of the OSLA as a brief instrument to identify abnormal levels of arousal associated with delirium in patients with acute hip fracture. First, we assessed the psychometric characteristics of the OSLA, using exploratory factor analysis. Second, we assessed the diagnostic performance of the OSLA for delirium detection, because acute onset altered level of arousal is considered a strong indicator of delirium.[4] Third, to explore the potential utility of the OSLA in detecting changes in delirium longitudinally, we assessed its ability to detect within-person fluctuations in delirium status and symptom severity over several test occasions.

2 | MATERIALS AND METHODS

2.1 | Design

This was a secondary analysis of data from a prospective cohort study in older adults with acute hip fracture with and without delirium.[17] Participants were assessed in the 24 hours prior to their surgery, repeatedly up to 14 days post-operatively, and at 3, 6, and 12 months post-operatively. Data up to day 14 are reported here. The study was approved by the Scotland A Research Ethics Committee and written consent from patients or legal proxies was obtained.

2.2 | Participants

A total of 108 community dwelling patients were recruited from orthopaedic wards at the Royal Infirmary of Edinburgh, Scotland. Patients were eligible if they were aged over 60 years and had an acute hip fracture and spinal anaesthesia. Patients were not eligible if they were nursing home residents; had taken oral or inhaled steroids in the last 10 weeks; had significant Parkinson's disease or other comorbid diseases with a prognosis of less than 1 year; or had major communication difficulties such as aphasia or where English was not their first language.

2.3 | Measurements and procedures

The diagnosis of delirium was made by a geriatrician (RJH), aided by the use of the Confusion Assessment Method (CAM)[18] and Delirium Rating Scale-Revised-98(DRS-R98),[19] and supplemented with assessments of level of consciousness (Richmond Agitation-Sedation Scale (RASS)[14] and cognition (Mini-Mental State Examination).[20] Delirium was considered present when CAM scores were positive or the total DRS-R98 score was over 17.75. Assessments took place preoperatively, daily from post-operative days 1-4, on day 7 and once between days 10-14 or until transfer to a rehabilitation unit or discharge from hospital. Participants were assessed as frequently as possible, including once per weekend, although this was dependent on researcher availability and with the aim of not becoming burdensome for participants. Illness severity and comorbidity were measured using the Acute Physiology Age and Chronic Health Evaluation (APACHE) II score,[21] the Charlson Comorbidity Index (CCI)[22] and the number of regular medications taken on admission. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) was used to assess pre-existing cognitive impairment.[23]

The OSLA comprises four items, each assessing a different feature of arousal: eye opening, eye contact, posture, and movement (Table 2). The items were developed by R. J. H. and A. M. J. M. through a two-stage process. The first stage involved generating potential items through informal observations in routine clinical practice and reviewing items in existing arousal scales. The goal of this stage was to capture and operationalise the judgements that clinicians

make regarding level of arousal in practice. The second stage involved editing the list of potential items to construct a scale which would be: (a) rapid and simple to score, (b) based on observation after a brief encounter alone, (c) operationalised using brief behavioural descriptions, and (d) allow grading of severity of the arousal disturbance for each item. The four items have different numbers of grades of severity based on what grades could readily be distinguished on brief observation. A total score (range 0-15) is obtained by summation of scores on each of the four items, with higher scores indicating greater abnormality. The OSLA is scored after a brief interaction with the patient and based on behavioural observations rather than cognitive testing. It generally takes under 1 minute to complete and does not require a verbal response from the patient.

All tests were administered by R. J. H. in the same order at the participants' hospital bedside.

2.4 | Statistical analysis

Analysis was carried out using R version 3.0.1.[24] Cases were excluded pairwise where data were missing. A threshold of $P < .05$ was taken to denote statistical significance.

Demographics and test scores for each group are presented as medians (interquartile range [IQR]) unless otherwise specified. Comparisons of OSLA scores and all other data were made between groups with and without delirium using Mann-Whitney *U* tests with continuity correction, separately for each assessment and also pooled assessments. Estimates of effect size *r* were calculated by dividing *z*-scores by the square-root of *n*.[25] Pearson's chi-squared tests were used for categorical data where appropriate. Correlations were calculated using Kendall's Tau due to frequent ties in the data. Holm corrections were applied to multiple comparisons.

We assessed the suitability of the data for exploratory factor analysis (EFA) using Bartlett's test of sphericity, the determinant of the correlation matrix and the Kaiser-Meyer-Olkin measure of sampling adequacy. Horn's parallel analysis (with 10 000 iterations) was used to empirically determine the number of factors to retain. Factors with eigenvalues greater >1 were assumed to be meaningful. We then conducted minimum residuals EFA without rotation—as parallel analysis suggested only a single factor—and used factor loadings of 0.40 or greater in the factor designation.

OSLA scores were compared between groups with and without delirium at each assessment. A receiver operating characteristic (ROC) analysis was conducted on OSLA scores with delirium diagnosis as a reference to assess the ability of the OSLA to detect delirium, for data collapsed across assessments and also separately for each assessment. The relationship between scores on the OSLA and the DRS-R98 severity scale was examined using Kendal-Tau correlations.

Exploratory linear mixed effects models were used to evaluate responsiveness of the OSLA to change in delirium status and severity over time, to provide additional information on its performance and also to explore the importance of level of arousal in contributing to a delirium diagnosis. These analyses fitted a within-person fluctuation model to test if changes in OSLA over time (of the order of days) were predicted by changes in each of the two time-varying predictors. Two models were fitted, one with delirium *diagnosis* as time-varying predictor (mean-centred) and all covariates (model 1), and one with delirium *severity* as time-varying predictor and all covariates (model 2) (R function lmer[24]). The models included the following time-invariant covariates: age, sex, IQCODE, CCI, and APACHE II score. The dependent variable was OSLA score.

All time-varying variables were standardised into units of the SD at baseline (the pre-operation assessment) with all means measured from the baseline mean. Continuous covariates were centred on the respective sample median: age (82 years), CCI (score of 1) and APACHE II (score of 8). Sex and IQCODE (scores ≥3.44 indicating dementia) were represented by a dummy variable. All models included the main effects of the covariates and their interaction with time.

The two time-varying predictors were decomposed into two variables: a person-mean (PM) variable and a within-person (WP) centred variable. In both models the WP variable entered the model at level-1, and the PM variable at level-2. Here we only report results of the WP variables. The 'WPseverity' effect represents how variation in OSLA over time fluctuates in step with variation in delirium severity. Likewise, the 'WPdiagnosis' effect represents how variation in OSLA over time fluctuates in step with delirium diagnosis.

3 | RESULTS

3.1 | Study sample

Participants were aged 61 to 95 (median = 82, IQR = 75-87) years. Baseline and demographic characteristics are provided in Table 3. There was some attrition of participants during the perioperative period, with seven participant withdrawals, three exclusions (two due to non-operative management of the fracture, and one due to an unpredictable and significant complication) and one death. The number of patients who provided data for the OSLA varied across assessments: pre-operative = 108, day 1 = 96, day 2 = 86, day 3 = 47, day 4 = 98, day 7 = 77, day 14 = 48 (reasons of missing values are presented in Table S1). The overall rate of delirium was 40.7% of patients.

3.1.1 | Psychometric characteristics of the OSLA

OSLA scores ranged between 0 and 9 (median = 2). Individual item score ranges were 0 to 3 (Eye Opening), 0 to 3 (Eye Contact), 0 to 2 (Posture), and 0 to 3 (Movement).

The four items of the OSLA were suitable for EFA, as indicated by Bartlett's test ($X^2[6] = 166.4$, $P < .001$), the determinant of the correlation matrix (0.74) and the overall (0.67) and item (all 0.67) Kaiser-Meyer-Olkin measure of sampling adequacy. A single factor, as indicated by a scree plot (Figure S1) and parallel analysis (Adjusted Eigenvalue = 1.76, Unadjusted Eigenvalue = 1.91, Estimated bias = 0.15)

accounted for 31% of the variance using minimum residuals factor analysis (Table S2).

3.1.2 | Association between OSLA and delirium status and severity

Scores on the OSLA were consistently higher in patients with delirium than those without delirium at each assessment, and also when data were collapsed across assessments (Table 4). A delirium diagnosis was always accompanied by an OSLA score greater than 0.

Further, scores on the OSLA correlated significantly with scores on the DRS-R98 severity sub-scale at each assessment point and when data were collapsed across assessments (Table 5). Thus, higher OSLA scores reflecting a greater degree of arousal abnormalities were associated with higher delirium symptom severity.

The Area under the ROC Curve for the OSLA for detecting delirium when data were pooled across assessments was 0.82 ($P < .001$, 95% CI [0.77, 0.86]). The Area under the ROC Curve analyses for data from individual assessment days yielded similar results to those of the pooled analysis (Table S3). Using an OSLA cutoff score ≥ 2, sensitivity and specificity for delirium were 0.87 (95% CI [0.84, 0.93]) and 0.53 (95% CI [0.48, 0.58]), respectively. A higher cutoff score of ≥3 was associated with a decline in sensitivity to 0.65 (95% CI [0.56, 0.74]) with an increased specificity of 0.85 (95% CI [0.81, 0.88]). The previously suggested OSLA cutoff score of ≥4 for delirium detection[4] resulted in a sensitivity of 0.42 (95% CI [0.33, 0.51]) and specificity of 0.95 (95% CI [0.93, 0.97]).

3.1.3 | Utility of the OSLA for measuring longitudinal changes in level of arousal associated with delirium

A change in diagnosis from delirium to no delirium was associated with a within-person decrease in the patient's OSLA score of 3.09 (Model 1: $\beta = -3.09$, SE = 1.41, $P < .03$). Further, a within-person increase in DRS-R98 delirium severity score of 1 unit (SD) relative to

TABLE 3 Patient demographic characteristics according to the presence of delirium at a minimum of one assessment point or the absence of delirium at all assessment points during hospital stay

	Delirium	No delirium	Comparison
Male:Female ratio (% Female)	20:24 (36)	22:42 (64)	$P = .25$, 95% CI [0.67, 3.75], $X^2(1) = 1.35$
Age (years)	83 (77-88)	81 (71-86)	$P = .163$, 95% CI [−6, 1], $U = 1185$, $z = -1.4$
n	44	64	
Length of hospital stay (days)	62 (28-81)	16 (11-30)	$P < .001$, 95% CI [17, 48], $U = 529$, $z = -4.6$
n	49	59	
Charlson Comorbidity Index	1 (0–2)	1 (0-1)	$P = .064$, 95% CI [−0, 0], $U = 1124$, $z = -1.9$
n	44	64	
Frailty Index on admission	0 (0–2)	0 (0-0)	$P = .007$, 95% CI [0, 1], $U = 1045$, $z = -2.7$
n	44	64	
IQCODE	3 (3-3)	3 (3-4)	$P < .001$, 95% CI [0, 0], $U = 666$, $z = -3.7$
n	40	57	
APACHE II	9 (7-11)	8 (6-10)	$P = .007$, 95% CI [−2, 0], $U = 977$, $z = -2.7$
n	44	64	

Note: Age, Acute Physiology and Chronic Health Evaluation (APACHE) II score range 0 to 71, higher score indicating greater illness severity. Charlson Comorbidity Index score range 0 to 31, higher score indicating greater comorbidity. Frailty index range 0 to 3, higher score indicating greater frailty. Informant Questionnaire of Cognitive Decline in the Elderly (IQCODE) average score range 1 to 5, higher score indicating greater cognitive decline. Descriptive statistics for continuous variables are expressed as medians (interquartile range). Ratios for categorical variables are expressed as frequencies (%).

TABLE 4 Results of the Observational Scale of Level of Arousal according to the presence or absence of delirium

	Pre-operative	Day 1	Day 2	Day 3	Day 4	Day 7	Day 14	Overall
Delirium	3 (2–4)	4 (2-5)	4 (2-4)	3 (3-5)	3 (2-4)	3 (3-6)	4 (4-6)	3 (2-5)
n	13	26	17	14	22	8	3	103
No delirium	2 (1–2)	2 (1-3)	2 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)	1 (0-2)	1 (1-2)
n	94	70	69	34	75	68	44	454
Comparison	$P < .001$, 95% CI [1, 2], $r = 0.36$	$P < .001$ 95% CI [1, 3], $r = 0.46$	$P < .001$ 95% CI [1, 2], $r = 0.37$	$P < .001$ 95% CI [1, 3], $r = 0.54$	$P < .001$ 95% CI [1, 2], $r = 0.50$	$P < .001$ 95% CI [1, 4], $r = 0.39$	$P = .005$ 95% CI [1, 6], $r = 0.41$	$P < .001$ 95% CI [1, 2], $r = 0.44$

TABLE 5 Correlation of DRS-R98 severity scores with the Observational Scale of Level of Arousal

	Pre-operative	Day 1	Day 2	Day 3	Day 4	Day 7	Day 14	Overall
Correlation coefficient	0.35	0.45	0.46	0.44	0.50	0.47	0.58	0.46
n	107	90	81	39	94	76	46	533
P-value, 95% Confidence Intervals (CI)	*P* < .001, 95% CI [0.17, 0.50]	*P* < .001, 95% CI [0.26, 0.60]	*P* < .001, 95% CI [0.26, 0.61]	*P* < .001, 95% CI [0.14, 0.44]	*P* < .001, 95% CI [0.33, 0.64]	*P* < .001, 95% CI [0.27, 0.63]	*P* < .001, 95% CI [0.35, 0.74]	*P* < .001, 95% CI [0.39, 0.46]

*Note:*DRS-R98, Delirium Rating Scale-Revised-98. Calculated using Kendall's Tau with Holm correction.

its average was associated with an increase in OSLA score of 0.5 (Model 2: ß = 0.50, SE = 0.13, *P* < .001).

4 | DISCUSSION

The present study provides preliminary support for the utility of the OSLA as a brief, accurate instrument for measuring level of arousal in delirium. The OSLA showed gradations in scores and was responsive to a change in delirium diagnosis within patients over time. Factor analysis of the eye opening, eye contact, movement, and posture items confirmed that it is appropriate to group these four features under a single factor (ie, arousal).

A wide range of OSLA scores was seen in the present study (0-9) and a delirium diagnosis was always accompanied by an OSLA score greater than 0. The OSLA therefore appears to be sensitive to the gradations in level of arousal seen in delirium in this non-ICU patient population.

Importantly, the OSLA was responsive to fluctuations in delirium status and severity over time within individual patients. This further supports the utility of the OSLA for assessing an aspect of delirium severity, and provides initial support for the utility of the OSLA in monitoring change in arousal as part of delirium in patients over time.

Chester et al[7] report a similar sensitivity for the mRASS (0.64) compared to the OSLA (0.65 using a cutoff ≥3) but with higher specificity (mRASS: 0.93, OSLA: 0.85). The latter finding may partly have resulted from incorporating attention assessments into the mRASS. Interestingly, the OSLA in our study performed broadly comparably while assessing level of arousal alone, without any items specifically assessing attention. Of note, the present findings suggest a lower cutoff point of ≥2 or ≥ 3 compared to the previously suggested optimal cutoff of ≥4 for the OSLA.[4]

Scores on the OSLA and the DRS-R98 severity scores were associated at each assessment point, even though the latter does not measure level of arousal explicitly. This likely reflects the hierarchical relationship between arousal and cognition, whereby level of arousal must be sufficient before cognition can be reasonably tested. This finding suggests that level of arousal may provide a useful, practical marker for grading severity of delirium.

The OSLA could have utility in operationalising the subtler arousal changes that appear to indicate delirium in general hospital populations. Specifically, the OSLA allows arousal features to be scored independently of one another, enabling observers to characterise a patient's level of arousal in some detail, while retaining its brevity and providing a measure of arousal in a single severity score. As such, it might prove useful as a brief, standardised delirium classification method (ie, hypo- vs hyperactive delirium) based on level of arousal alone.[15] Although the OSLA was developed as a stand-alone test, it could complement existing delirium assessment batteries. Of note, a brief combined arousal-attention assessment using OSLA and SAVEAHAART has been shown to have high diagnostic accuracy for detecting delirium even in a subgroup of patients with dementia, and thus could have clinical utility for diagnosing delirium superimposed on dementia.[16,26]

This study has several limitations. Delirium diagnosis and arousal assessments were done by a single rater, hence estimations of interrater reliability were not possible, and the diagnostic accuracy of the OSLA for delirium may have been inflated. Formal assessments of interrater reliability and validity using independent, blinded raters are required to supplement the current findings. Further, a number of patients were lost to follow-up in the first 14 days post-surgery. In the pre-fracture and immediate post-operative stages, scoring of the Movement item may have been confounded by the patient's fracture, which could restrict movement. Also, the different numbers of grading levels for each item might mean that some items carry greater weight on the scale.

Future studies are needed to evaluate the utility of OSLA in different populations (eg, palliative care, emergency departments) and to assess the prognostic value of higher OSLA scores for unfavourable outcomes in prospective cohort studies. A recent study provided preliminary support for the utility of the OSLA as a tool for detailed measurement of level of arousal in ICU patients,[27] though the OSLA itself has yet to be formally validated in this population.

In conclusion, this study provides promising evidence in support of the OSLA as a method for arousal assessment in the context of delirium. The OSLA may usefully complement existing measures of delirium where additional detail is desirable.

ACKNOWLEDGEMENTS

This work was supported by a Clinical Research Training Fellowship to R. J. H. from Research into Ageing (AgeUK) and the British Geriatrics Society (grant number 342). Funding from the Biotechnology and

Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, the Economic and Social Research Council, and the Medical Research Council is gratefully acknowledged.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available onrequest from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Zoë Tieges https://orcid.org/0000-0002-3820-3917

REFERENCES

1. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383(9920):911-922.
2. Reynish EL, Hapca SM, De Souza N, Cvoro V, Donnan PT, Guthrie B. Epidemiology and outcomes of people with dementia, delirium, and unspecified cognitive impairment in the general hospital: prospective cohort study of 10,014 admissions. *BMC Med*. 2017;15(1):140.
3. Marcantonio ER. Delirium in hospitalized older adults. *New Engl J Med*. 2017;377(15):1456-1466.
4. Tieges Z, McGrath A, Hall RJ, Maclullich AM. Abnormal level of arousal as a predictor of delirium and inattention: an exploratory study. *Am J Geriatr Psychiatry*. 2013;21(12):1244-1253.
5. Morandi A, Davis D, Bellelli G, et al. The diagnosis of delirium superimposed on dementia: an emerging challenge. *J Am Med Dir Assoc*. 2017;18(1):12-18.
6. Bellelli G, Morandi A, Davis DH, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing*. 2014;43(4):496-502.
7. Chester JG, Beth Harrington M, Rudolph JL, Group VADW. Serial administration of a modified Richmond Agitation and Sedation Scale for delirium screening. *J Hosp Med*. 2012;7(5):450-453.
8. Bleyer AJ, Vidya S, Russell GB, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. *Resuscitation*. 2011;82(11):1387-1392.
9. Royal College of Physicians. *National Early Warning Score (NEWS): Standardising the assessment of acute illness severity in the NHS*. Report of a working party. London; 2017.
10. Brown LJ, Fordyce C, Zaghdani H, Starr JM, MacLullich AM. Detecting deficits of sustained visual attention in delirium. *J Neurol Neurosurg Psychiatry*. 2011;82(12):1334-1340.
11. Lowery DP, Wesnes K, Brewster N, Ballard C. Quantifying the association between computerised measures of attention and confusion assessment method defined delirium: a prospective study of older orthopaedic surgical patients, free of dementia. *Int J Geriatr Psychiatry*. 2008;23(12):1253-1260.
12. Tieges Z, Brown LJ, Maclullich AM. Objective assessment of attention in delirium: a narrative review. *Int J Geriatr Psychiatry*. 2014;29(12):1185-1197.
13. Tieges Z, Evans JJ, Neufeld KJ, MacLullich AMJ. The neuropsychology of delirium: advancing the science of delirium assessment. *Int J Geriatr Psychiatry*. 2018;33(11):1501-1511.
14. Sessler CN, Gosnell MS, Grap MJ, et al. The Richmond Agitation-Sedation Scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med*. 2002;166(10):1338-1344.
15. Han JH, Brummel NE, Chandrasekhar R, et al. Exploring Delirium's heterogeneity: association between arousal subtypes at initial presentation and 6-month mortality in older emergency department patients. *Am J Geriatr Psychiatry*. 2017;25(3):233-242.
16. Richardson SJ, Davis DHJ, Bellelli G, et al. Detecting delirium superimposed on dementia: diagnostic accuracy of a simple combined arousal and attention testing procedure. *Int Psychogeriatr*. 2017;29(10):1585-1593.
17. Hall RJ, Ferguson KJ, Andrews M, et al. Delirium and cerebrospinal fluid S100B in hip fracture patients: a preliminary study. *Am J Geriatr Psychiatry*. 2013;21(12):1239-1243.
18. 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(12):941-948.
19. Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N. Validation of the Delirium Rating Scale-revised-98: comparison with the delirium rating scale and the cognitive test for delirium. *J Neuropsychiatry Clin Neurosci*. 2001;13(2):229-242.
20. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
21. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13(10):818-829.
22. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383.
23. Jorm AF. A short form of the informant questionnaire on cognitive decline in the elderly (IQCODE): development and cross-validation. *Psychol Med*. 1994;24(1):145-153.
24. *R: a language and environment for statistical computing* [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2014.
25. Field A, Miles J, Field Z. *Discovering Statistics Using R*. London, England: Sage; 2012.
26. Quispel-Aggenbach DWP, Holtman GA, Zwartjes H, Zuidema SU, Luijendijk HJ. Attention, arousal and other rapid bedside screening instruments for delirium in older patients: a systematic review of test accuracy studies. *Age Ageing*. 2018;47(5):644-653.
27. Tang E, Laverty M, Weir A, et al. Development and feasibility of a smartphone-based test for the objective detection and monitoring of attention impairments in delirium in the ICU. *J Crit Care*. 2018;48:104-111.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Hall R, Stíobhairt A, Allerhand M, MacLullich AMJ, Tieges Z. The Observational Scale of Level of Arousal: A brief tool for assessing and monitoring level of arousal in patients with delirium outside the ICU. *Int J Geriatr Psychiatry*. 2020;1–7. https://doi.org/10.1002/gps.5324