Sedation / Delirium

Development and feasibility of a smartphone-based test for the objective detection and monitoring of attention impairments in delirium in the ICU

Elaine Tang [a,b,1], Mairi Laverty [a,b,1], Alexander Weir [c], Elizabeth S. Wilson [d], Timothy S. Walsh [e], Mike Allerhand [f], Stuart Parks [c], Alasdair M. J. MacLullich [b,f,2], Zoë Tieges [b,f,*,2]

[a] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK
[b] Edinburgh Delirium Research Group, Royal Infirmary of Edinburgh, University of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK
[c] Medical Devices Unit, West Glasgow Ambulatory Care Hospital, Dalnair Street, Glasgow G3 8SJ, UK
[d] Critical Care Medicine and Anaesthesia, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK
[e] Department of Anaesthesia Critical Care and Pain Medicine, Royal Infirmary of Edinburgh, University of Edinburgh, 51 Little France Crescent, Edinburgh EH16 4SA, UK
[f] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

ARTICLE INFO

ABSTRACT

Purpose: Delirium in the ICU is under-diagnosed. We evaluated feasibility and performance of a novel smartphone-based test for objectively detecting inattention in delirium.
Material and methods: DelApp-ICU combines a behavioural assessment and an attention task, whereby participants follow simple commands and count serially presented circles (score range 0–12, lower scores indicating worse performance). We assessed feasibility through staff interviews. Then we performed a preliminary case-control study in patients with and without delirium (ascertained with the Confusion Assessment Method for the ICU) who underwent the DelApp-ICU on up to 4 days.
Results: Forty-six patients (median age = 57.5 years, range 18–83) were assessed 89 times in total (N's = 46, 29, 10 and 4 for subsequent assessments; 33.7% delirious). DelApp-ICU scores were lower in delirium (N = 20; median = 0.5, Inter-Quartile Range (IQR) = 0–4.75) compared to no delirium (N = 26, median = 12, IQR = 8–12) on days 1, 2 and 3 (p < 0.001, p < 0.001 and p < 0.05, respectively). A DelApp-ICU score ≤6 was 100% sensitive and 96% specific to delirium on day 1. Positive and Negative Predictive Values were 91% and 100%, respectively. DelApp-ICU scores were responsive to changes in CAM-ICU status.
Conclusions: DelApp-ICU shows promise for assessing inattention and delirium in ICU patients, including longitudinally monitoring deficits and providing a metric of delirium severity.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Delirium is an acute, serious neuropsychiatric disorder which occurs in up to 80% of ventilated patients in the ICU [1]. Delirium in ICU patients is associated with adverse outcomes including longer hospitalisation, higher medical costs, increased mortality and greater risk of developing long-term cognitive impairment, particularly when the delirium has a prolonged duration [2-5]. Delirium detection requires accurate assessment of disturbances in level of arousal and attention, as well as determining the onset of these features. Further, using a longitudinal assessment method which has high validity and reliability and which is responsive to change is crucial for tracking the time course of delirium [6,7]. Timely diagnosis of delirium is vital to allow early intervention and effective management of precipitants and distressing symptoms common in delirium [8-10]. However, despite its clinical importance, delirium in the ICU often remains undetected [2].

The most commonly used tools used for the detection of ICU delirium are the Confusion Assessment Method for the ICU (CAM-ICU [1,11]) and the Intensive Care Delirium Screening Checklist (ICDSC [12]). Whereas the initial validation studies of these tools have shown high sensitivity, specificity and inter-rater reliability for the detection of delirium, others have reported that diagnostic accuracy is partly dependent on tester expertise [13,14]. A recent multi-centre study found that the CAM-ICU had excellent specificity and moderate sensitivity in routine practice [15].

We previously developed a novel computerized test of attention implemented on a custom-built hardware device termed the Edinburgh Delirium Test Box (EDTB), which successfully quantifies inattention using a counting up task paradigm. It has been found to differentiate

between general ward patients with and without delirium [16,17]. A pilot study using a modified paradigm (EDTB-ICU) suggested that this test also performs well in detecting delirium in ICU patients with RASS > −3. The EDTB-ICU also showed promising utility for longitudinally tracking attentional deficits [18].

Although the EDTB-ICU showed excellent performance, it is clear that tests on devices that are universally available, such as smartphones, would be more readily applicable in research and clinical settings. Therefore, we developed a software application for detecting deficits of visual sustained attention in delirium (DelApp) based on the EDTB paradigm. A preliminary study demonstrated initial feasibility of DelApp and excellent accuracy for discriminating delirium from dementia and cognitively unimpaired general ward patients [19].

More recently we have developed a modified DelApp assessment for use in ICU settings. The DelApp-ICU is based on the EDTB-ICU [18]; it includes a brief behavioural assessment and is suitable for use in non-verbal patients.

The aim of this preliminary study was to evaluate feasibility and performance of the DelApp-ICU as a novel instrument for detecting and monitoring both level of arousal and inattention in delirium in ICU patients. First, we evaluated feasibility of the DelApp-ICU through interviews with ICU staff, followed by a case-control study of the performance of the DelApp-ICU in patients with and without delirium, which included longitudinal assessments.

2. Material and methods

2.1. Ethics

The study was approved by the Scotland A Research Ethics Committee.

2.2. Main study

2.2.1. Participants

Patients admitted to the general ICU of the Royal Infirmary of Edinburgh were screened for eligibility. Exclusion criteria were age under 18 years, known pre-existing cognitive impairment, brain injury, visual or hearing impairments severe enough to preclude cognitive testing, and photosensitive epilepsy. Patients with a Richmond Agitation and Sedation Scale Score (RASS) of ≤ −3 (any movement but no eye contact, to voice) [20] were excluded so as to align with the criteria of a previous study [18], and because this was an exploratory study, we wanted to include as many ICU patients in our sample able to undergo the sustained counting task (Fig. 1). The sample size was determined based on a previous pilot study using the EDTB-ICU, where large effect sizes were found [18].

2.2.2. Measurements and procedures

Patients were assessed routinely for delirium with the CAM-ICU [11] and RASS [20] by clinical staff trained in using these instruments. The researcher recording the reference standard was not blinded to CAM-ICU status. Additionally, the Observational Scale of Level of Arousal (OSLA) was used by the researchers to provide a more fine-grained measure of level of arousal [16]. Higher scores on the OSLA indicate greater abnormality in level of arousal (range = 0–15). The Acute Physiology And Chronic Health Evaluation (APACHE II) score was derived from the case notes [21].

2.2.2.1. DelApp-ICU assessment.
The DelApp-ICU commenced with a behavioural assessment to determine the participant's level of arousal and basic orienting response (a component of attention), and to select a suitable means of non-verbal communication if required from the following possible response methods: pointing, thumbs up, squeezing the administrator's hand, nodding, or sticking one's tongue out. Next, the ability to track the assessor's moving finger for a period of 5 s was

determined. Finally, the participant's capacity to follow instructions was assessed by instructing them to identify a given number on a response card held in front of them by the assessor. The A4-sized laminated response card contained the numbers '3, 4, 5, 6'. The assessor would point to each of the numbers and ask the participant to indicate when the number corresponding to the correct answer was pointed out via verbal or non-verbal (as mentioned above) methods. The maximum possible score for this behavioural assessment was 3.

Next, a visual task was presented on the 9.2 × 5.9 cm display of a Samsung Galaxy S2 smartphone running the Android 3.2 operating system, held at a distance of approximately 30 cm from the participant's eyes. This task consisted of four short trials which required participants to (1) identify a change in colour of a circle (5 cm diameter) presented on the smartphone screen from white to grey or vice versa (trials 1 and 2) and (2) identify a change in the shape of the stimulus from circle to star or vice versa (trials 3 and 4). For both these trials, participants were asked to indicate verbally or non-verbally (using the method selected during the behavioural assessment) when they noticed a change in colour or shape.

Participants then proceeded with the visual sustained attention task, which was deemed to place relatively high demands on attentional functioning whilst being of low overall task complexity. Brightness was set automatically based on the ambient light. The target stimulus, a large white circle (5 cm diameter), was displayed in the centre of the smartphone screen against a black background and would be intermittently displayed (i.e. appear and disappear) a number of times throughout the trial. On some trials, the circle was surrounded by pseudo-random subsets of 8 small downward-pointing triangles (0.3 × 0.3 × 0.4 cm), each presented for 500 or 700 ms duration, to provide distraction (see Supplementary Material). Participants were instructed to count the number of times the circle was shown and to ignore any triangles which may appear in later trials. For each trial, when asked by the assessor how many circles were shown, participants responded either verbally, or by selecting an answer from four possible answers presented on the response card using non-verbal methods.

The counting task comprised a practice trial (which was not scored) and nine scored trials. On all trials, target lights were presented between 3 and 6 times per trial for a fixed duration of 1000 ms, with an inter-stimulus interval of 1000 ms, 2000 ms, or 3000 ms which varied across trials. The trials were presented in a fixed order. The practice trial consisted of a white circle which was displayed three times at a fixed inter-stimulus interval of 2000 ms with concurrent presentation of distracters (700 ms duration each). If the participant responded incorrectly, the practice trial was repeated once more before proceeding with the nine remaining trials. All responses were recorded and scored as correct or incorrect; if no response was given then this was scored as an error.

Total DelApp-ICU score was calculated as the summed score of the behavioural assessment (0–3) and the sustained attention test (0–9), yielding a total score ranging between 0 (severe attentional deficits) and 12 (no attentional deficits) (Fig. 2).

A pilot study in 10 non-delirious verbal ICU patients was conducted as the ethics committee ruled it unnecessary to involve patients requiring proxy consent at that stage. All participants obtained the maximum possible score of 3 for the behavioural assessment, and all completed the attention task with 6 out of 10 participants making no errors. Thus, the DelApp-ICU was found to be feasible for use in the ICU setting in non-intubated patients without delirium.

2.2.3. Assessment of staff acceptability

Twenty-two ICU staff were interviewed using a semi-structured format in order to gain insight into (i) perceptions on delirium assessment in routine clinical work; and (ii) opinions on the feasibility of using a smartphone application in the ICU for the assessment of delirium. The ICU staff interviews focused on general knowledge of delirium, opinions on the use of CAM-ICU for routine screening of delirium, and opinions
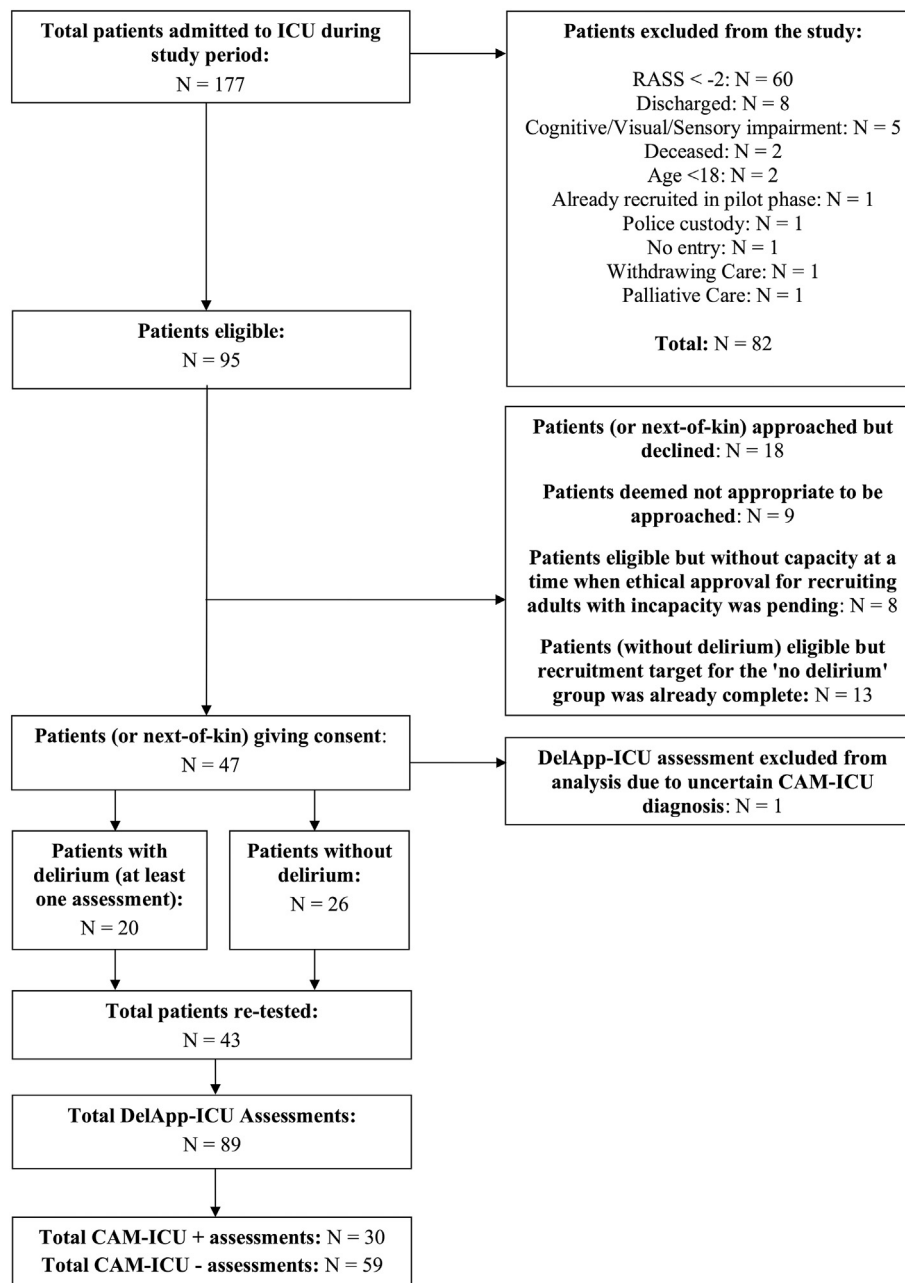
Fig. 1. Recruitment flowchart.

on the use of smartphone apps in general and specifically the DelApp-ICU for delirium assessment (see Supplementary Material). Nurses were shown the DelApp-ICU during the interview. Staff interviews were audio-recorded and transcribed, and responses were grouped according to pre-defined themes.

2.2.4. Longitudinal case-control study

This was a prospective case-control study of ICU patients who were recruited into groups with and without delirium at the time of study entry. All patients underwent the DelApp-ICU at the initial assessment and then underwent up to 3 further assessments on separate days, only if the patient remained in the ICU at the time of further assessments and if they, their next of kin and/or the healthcare team agreed to continued participation. The longitudinal design was used to evaluate the performance of the DelApp-ICU for quantifying fluctuations in attention impairments and delirium over time.

Five sets of the DelApp-ICU counting task were constructed to administer to patients undergoing longitudinal assessments, to minimize practice effects. Only the number of target illuminations, though all between 3 and 6 illuminations per trial, differed between these sets. All patients were assessed using set 1 on the first assessment day; sets 2–5 were administered in a randomised manner on subsequent occasions.

2.2.5. Statistical Analysis

Non-parametric statistical tests were used because the majority of the data was not normally distributed and there were unequal variances across groups. Comparisons of DelApp-ICU test scores and all other data were made between groups with and without delirium using Mann-Whitney U tests and Chi Square tests. Receiver Operating Curve (ROC) analyses were conducted on DelApp-ICU scores on the first assessment day.

Behavioural Assessment

Does the patient respond to verbal stimuli? → **No** → The patient cannot be assessed and scores 'N/A'. End of assessment.

↓ **Yes**

Is the patient able to track moving stimuli? → **No** → The patient scores '0' for this assessment. End of assessment.

↓ **Yes**

Is the patient able to follow simple commands?

- Verbally indicate a number from a response card
- Point to the response card
- Give a thumbs-up
- Squeeze the researcher's hand
- Nod
- Stick out tongue

→ **No** → The patient scores '1' for this assessment. End of assessment.

→ **Yes** → The patient scores '2' for this assessment if able to respond (regardless of accuracy) and proceeds with the visual task.

The patient scores '3' for this assessment and proceeds with the visual task.

Visual Task

Task 1: Can the patient see a change in colour?

Task 2: Can the patient see a change in shape?

→ 2 trials per task (total 4 trials).

If patient is able to see at least one change, they proceed with the counting task.

Note: scores not included in total DelApp-ICU score.

↓

DelApp-ICU Counting Task: Tests of Sustained Attention

Level 1: No distracters

Level 2: Medium distraction

Level 3: High distraction

→ **3 trials per level:**

Between 3-6 circles per trial.

Score: 1 (correct), 0 (incorrect), N/A (incomplete) for each trial.

Trial 1: 1000ms inter-stimulus gap

Trial 2: 2000ms inter-stimulus gap

Trial 3: 3000ms inter-stimulus gap

↓

Scoring:

Behavioural Assessment = ___/3
Counting task Level 1 = ___/3
Counting task Level 2 = ___/3
Counting task Level 3 = ___/3
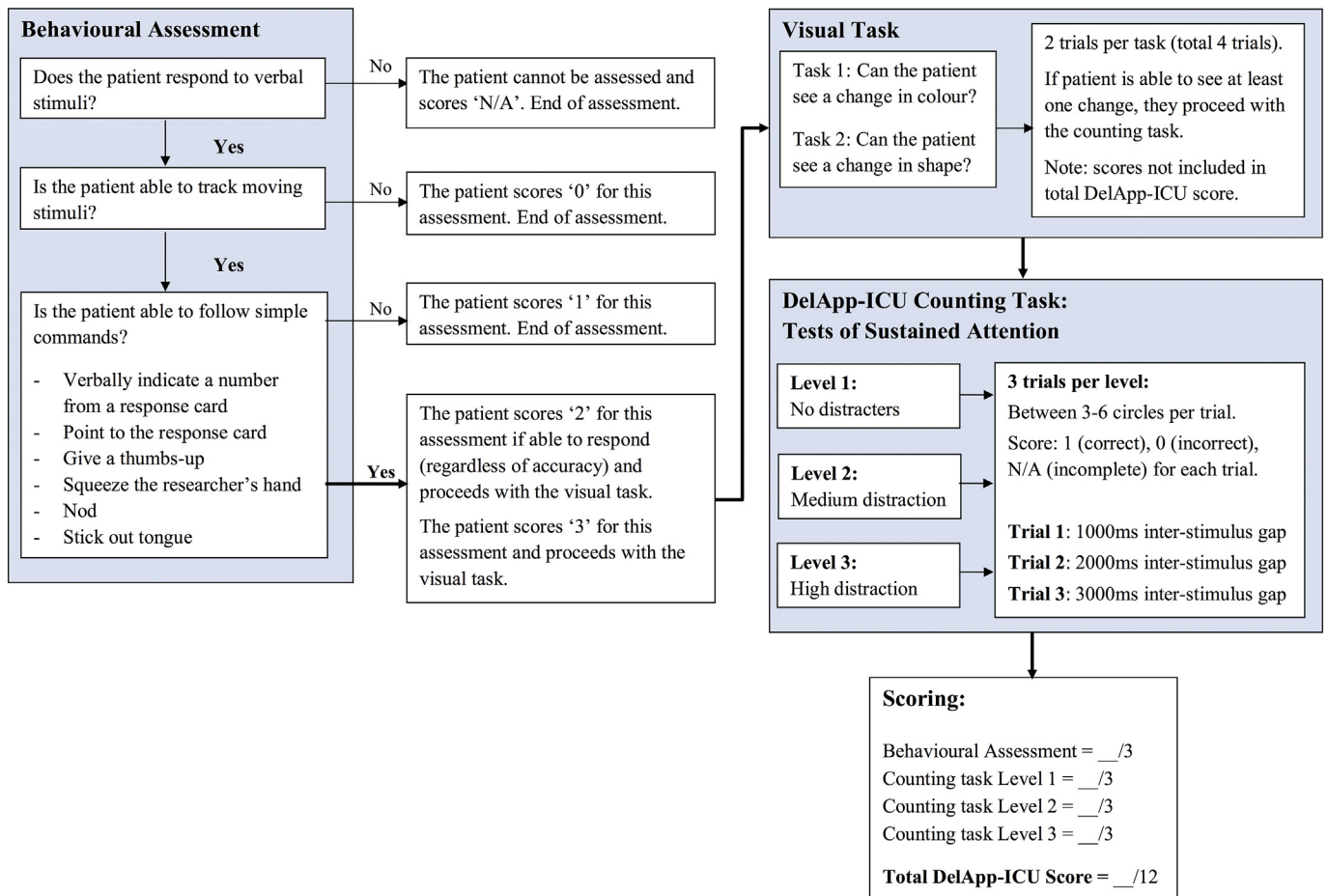
Total DelApp-ICU Score = ___/12

Fig. 2. Flow diagram of DelApp-ICU testing procedure. All participants undertook a behavioural assessment to examine their level of arousal and ability to engage with cognitive testing, as well as establishing a feasible method of response. Participants who successfully completed this assessment then undertook a brief sustained attention task implemented on a smartphone.

A linear mixed effects model (R function *lmer* [22]) was fitted to the repeated-measures data in order to test whether the association between the scores on the CAM-ICU and DelApp-ICU varied over time. The model included CAM-ICU score (coded 0 and 1), the assessment day (coded 0 to 3), and their interaction as predictors of the DelApp-ICU score. The time-varying association between scores on the CAM-ICU and DelApp-ICU was allowed to have a random intercept between patients.

Non-parametric tests (Mann-Whitney *U*) were used to test for differences between patients who completed one assessment and those who completed multiple assessments on a range of baseline variables. Analyses were carried out with R [22].

3. Results

3.1. Acceptability of DelApp-ICU to clinical staff: nurse interviews

Twenty-two ICU staff (20 nurses, one doctor and one nurse practitioner) were interviewed. The majority of nurses reported using the CAM-ICU once at the beginning of each shift, and would sometimes re-administer this tool if sedation was reduced or in case of other changes in patient status such as fluctuations in delirium status.

Staff opinions of the CAM-ICU were mixed. Most interviewees found the CAM-ICU to be a useful, straightforward tool for detecting delirium, but some voiced issues regarding its use in clinical practice. These included challenges with test administration due to physical weakness of patients, excessive agitation or drowsiness, hearing impairment, or not being fluent in English. The CAM-ICU was perceived as sometimes time-consuming and requiring a lot of explanation for some patients.

Many nurses felt that the use of the CAM-ICU, or any other delirium assessment tool, was not necessary because they were able to judge a patient's delirium status without using a tool.

There were varied opinions of the use of smartphone applications in the ICU, and specifically the DelApp-ICU. At the time of the study, the use of phones was discouraged on the ward and consequently nurses were uncomfortable using a phone in front of visitors. Infection control was raised as another issue, if tests running on smartphones were to be used between beds. Anonymity of patients was also a concern because of the embedded cameras in the phones (even if these were not being used in the test). It was suggested that there would need to be at least one smartphone in each ward and preferably one for every bed space to reduce risk of infection. One respondent noted that the use of smartphones could negatively impact on the interaction between patient and nurse.

There were concerns that older patients might not respond well to a test presented on a smartphone and that it could make delirious patients more agitated. It was thought to be more suitable for use in younger patients, who might prefer smartphone-based tests over more traditional methods of assessment. However, most nurses thought that a smartphone test was an innovative way of moving forward.

The DelApp-ICU was found to be user friendly, but issues were noted regarding its practical application. Poor vision was frequently mentioned as a potential issue, specifically glasses not being available or uncomfortable to wear, the presence of cataracts and lack of prolonged eye opening or eye contact in many patients due to abnormal arousal and/or attention. Some interviewees suggested that a tablet device may be more appropriate as this would allow presentation of larger stimuli.

3.2. Longitudinal case-control study

3.2.1. Participants

Forty-seven patients were assessed during the main study phase in a period of 40 testing days. Twenty-one patients were diagnosed with delirium according to the CAM-ICU on at least one assessment, and 25 patients were never delirious during the study (Table 1). RASS scores ranged from −2 to +3 (median score = 0). Data from one patient were excluded due to uncertain clinical delirium diagnosis. A total of 30 delirious episodes and 59 non-delirious episodes were analysed. Twelve assessments were performed on ventilated patients. Of the 43 repeat assessments, 30 were performed the day after the previous assessment, 7 assessments were performed after two days, 3 assessments after three days, and 1 assessment each after 4, 6 and 9 days (median = 1 day, IQR = 1–2).

The DelApp-ICU assessment took approximately 3–8 min to administer, depending on the level of arousal and responses of the participant. Verbal responses were used in 72% of assessments, while pointing was the most common non-verbal method of response (7%) followed by hand squeezing (3%), although it was noted that verbal participants sometimes used non-verbal responses such as nodding and pointing in addition to speech. One ventilated participant used a combination of responses (nodding, pointing, squeezing and raising eyebrows). The remaining 18% of participants failed to pass the initial behavioural assessment, therefore did not proceed with the sustained attention

task. Importantly, all ICU patients recruited were able to use the DelApp-ICU, demonstrating that it is a feasible tool.

3.3. DelApp-ICU performance

Patients with delirium (N = 20) scored significantly lower on the DelApp-ICU task (median = 0.5, IQR = 0–4.75, range = 0–6) compared to patients without delirium (N = 26, median = 12, IQR = 8–12, range = 5–12) on the first assessment day (U = 0.5, p < 0.001) (Fig. 3). This difference was also evident for the second assessment (delirium: N = 8, median = 0, IQR = 0–3.75, range = 0–6; no delirium: N = 21, median = 11, IQR = 8–12, range = 4–12; U = 3.5, p < 0.001) and third assessment (delirium: N = 2, median = 0.5, IQR = 0–0, range = 0–1; no delirium: N = 8, median = 10.5, IQR = 8.25–11.75, range = 4–12; U = 0.00, p < 0.05). No patients were diagnosed with delirium at the fourth assessment, precluding group analysis.

ROC analysis of data from the first assessment showed that DelApp-ICU score was successfully able to differentiate between delirious and non-delirious patients (AUC = 0.999 (95% CI = 0.995–1.000), p < 0.001). A score of 6 or less was 100% sensitive and 96% specific to delirium on the first assessment, whereas a score of 5 or less was 95% sensitive and 100% specific to delirium. The Positive Predictive Value was 91% and Negative Predictive Value was 100% on first assessment for a cut off score of 6 or less.

The linear mixed-effects model showed that a diagnosis of delirium (according to the CAM-ICU) was associated with a decrease in the patient's DelApp-ICU score of −7.46 at the time of diagnosis (beta = −7.46, S.E. = 0.64, t (df = 82) = −11.71, p < 0.001) (Fig. 4). There was no significant interaction between the CAM-ICU effect and the assessment day (beta = 0.01, S.E. = 0.66, t (df = 51) = 0.02, ns), indicating that the association between CAM-ICU and DelApp-ICU scores did not change after the first assessment.

There were no differences in age, RASS, OSLA and APACHE II scores between those who completed only one assessment compared to those who underwent multiple assessments; hence, selection bias was deemed unlikely.

4. Discussion

This preliminary study demonstrated initial feasibility and validity of a smartphone test for detecting and monitoring attention and arousal
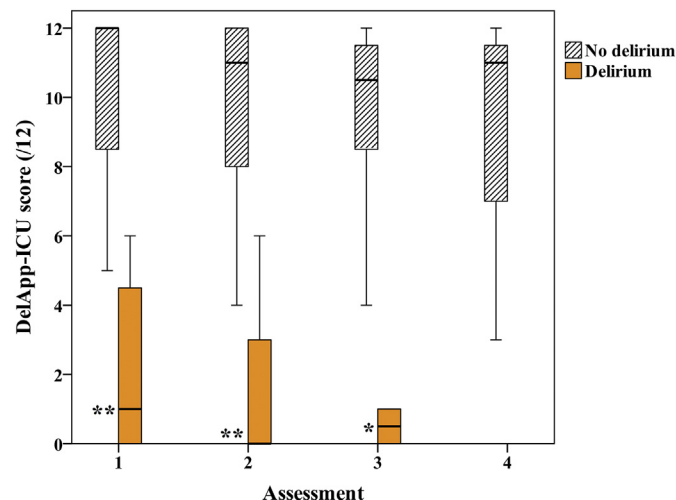
Table 1
Demographic characteristics of participants in the main study and characteristics of participants in the main study at each assessment.

	All patients	CAM-ICU positive on at least one assessment	CAM-ICU negative on all assessments
N	46	21	25
Age, years (median, IQR)	57.5 (49–70)	61 (48–77.5)	54 (46.5–68.5)
Sex (female (N, %))	17 (37.0%)	8 (38.1%)	9 (36.0%)
APACHE II Score	14 (12−20)	17 (13–24.5)	13 (10–15)
ICU length of stay, days (median, IQR)	4.9 (3.0–9.9)	6.8 (3.4–13.9)	4.3 (2.5–8.1)
ICU mortality (N, %)	1 (2.2%)	1 (4.8%)	0 (0%)

	All EDTB-ICU assessments	CAM-ICU positive assessments	CAM-ICU negative assessments
N	89	30	59
Mechanically ventilated N, %)	12 (13%)	8 (26.7%)	4 (6.7%)
RASS score (N, %)			
+3	1 (1.1%)	1 (3.3%)	0 (0.0%)
+2	0 (0.0%)	0 (0.0%)	0 (0.0%)
+1	6 (6.7%)	6 (20.0%)	0 (0.0%)
0	**64 (71.9%)**	**13 (43.3%)**	**51 (86.4%)**
−1	16 (18.0%)	8 (26.7%)	8 (13.6%)
−2	2 (2.2%)	2 (6.7%)	0 (0.0%)
OSLA score (N, %)			
0	**43 (48.3%)**	**2 (6.7%)**	**41 (69.5%)**
1–3	12 (13.5%)	3 (10.0%)	9 (15.3%)
4–6	13 (14.6%)	6 (20.0%)	7 (11.9%)
7–9	10 (11.2%)	8 (26.6%)	2 (3.4%)
10–12	10 (11.2%)	10 (33.3%)	0 (0.0%)
13–15	1 (1.1%)	1 (3.3%)	0 (0.0%)
Total DelApp-ICU score (median, IQR) (/12)	8 (3−12)	0 (0–4.25)	11 (8–12)
Behavioural assessment subscore (/3)	3 (3–3)	0.5 (0–3)	3 (3–3)
Attention assessment subscore (/9)	5 (0.75–9)	0 (0–1.25)	8 (5–9)

RASS = Richmond Agitation Sedation Scale. OSLA = Observational Scale for Level of Arousal, range 0–15 with higher scores indicating a greater degree of abnormal level of arousal (either decreased or increased).

Fig. 3. Boxplot illustrating the results of the DelApp-ICU test for patients with and without delirium according to CAM-ICU score at 4 assessment days. The median is represented by the thick horizontal bars. The interquartile range is represented by the height of the inner boxes. The upper and lower whiskers represent scores outside the middle 50%. *p < 0.05; ** < 0.001.
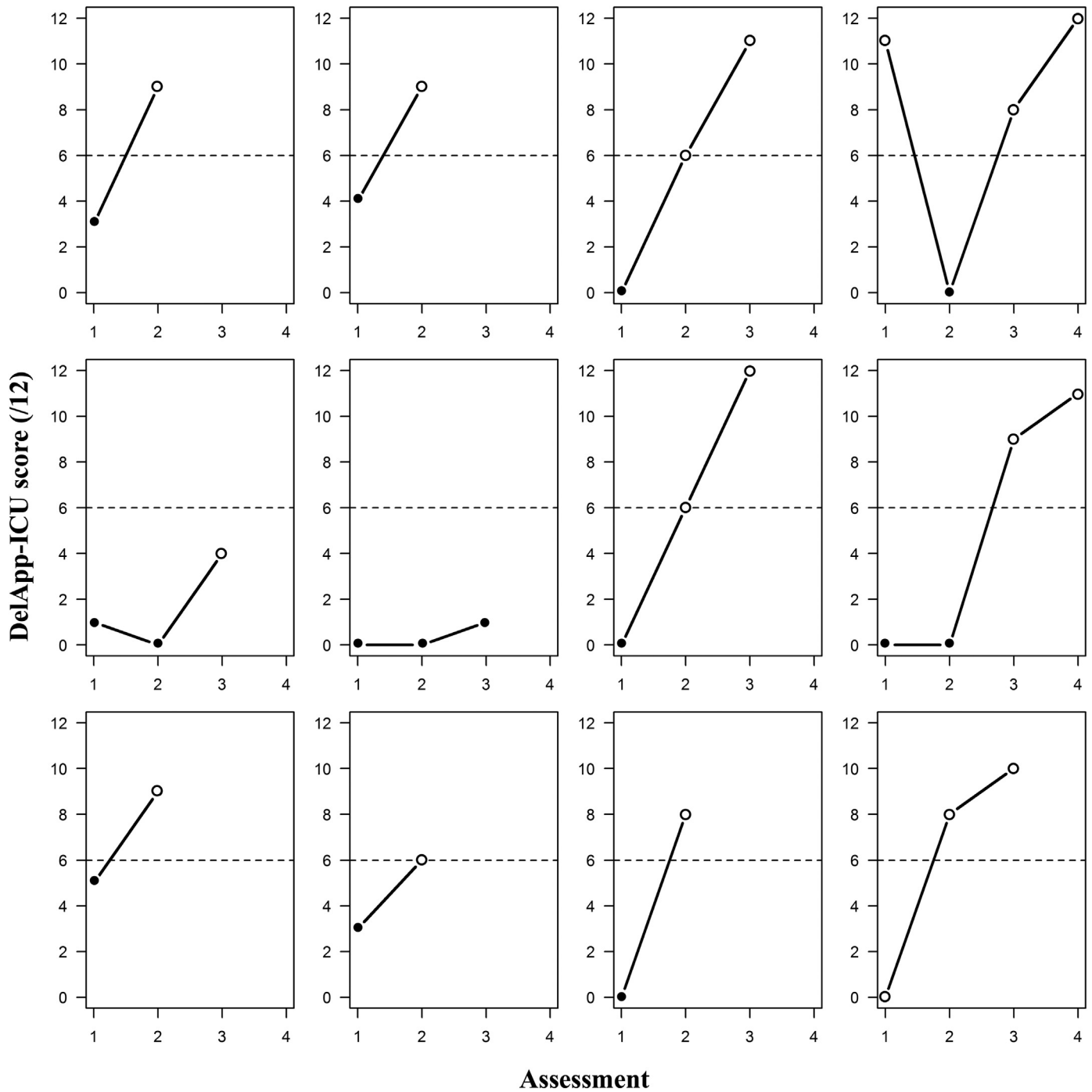
Fig. 4. DelApp-ICU scores across assessments for patients who were assessed when both CAM-ICU positive (black circles) and CAM-ICU negative (white circles). DelApp-ICU scores fluctuate with each assessment depending on delirious state.

impairments in delirium in the ICU. In the present exploratory work, the DelApp-ICU had excellent diagnostic performance in a selected sample of ICU patients, consistent with previous findings using a hardware version of the test [18]. DelApp-ICU scores were responsive to a within-person change in delirium diagnosis according to the CAM-ICU. The DelApp-ICU assessment was found to be suitable for use in intubated or otherwise non-verbal patients with varying levels of arousal.

4.1. Strengths and limitations

The DelApp-ICU test appears to have a number of strengths including its ease of use, objectivity and portability. It provides a standardised

assessment of attention impairment in delirium and it requires little training. The DelApp-ICU may be particularly suitable when used by non-experts as it does not rely strongly on clinical judgment, reducing assessor bias or error [23].

Differing from most existing diagnostic tools whereby attentional function is scored in a binary manner, the DelApp-ICU provides a graded measure of attention reflecting the continuum of attentional impairment seen in delirium. Indeed, DelApp-ICU scores varied across the full range of possible scores (0−12) with considerable variation noted even among patients with delirium (0–6). As such, the DelApp-ICU has potential utility as a valid measure of delirium severity, and may be useful in tracking change in presence and severity of attention

impairments in delirium over time. There is some agreement that quantifying severity of delirium in ICU is useful for detecting delirium at an early (possibly prodromal) stage, to assess whether delirium is persistent or resolving over time and consequently inform sedation practice, and to provide a means of measuring response to treatment [24]. Furthermore, early detection of delirium can minimize the risk of developing any adverse outcomes associated with delirium, including prolonged hospitalisation and increased mortality.

A common issue in assessing hospitalised patients for delirium is that impaired level of arousal often precludes formal cognitive assessment of its cardinal feature of impaired attention, especially in the most vulnerable patients. The DelApp-ICU addresses this issue of testability by including a brief assessment of arousal (as part of the behavioural assessment) and basic orienting, which is reflected in the total DelApp-ICU score. In other words, all patients including those who have difficulty engaging with the task due to drowsiness or extreme agitation are assessable with the DelApp-ICU process.

The study has a number of limitations. A number of potential barriers were identified from the nurse interviews, including concerns about patient anonymity, infection control, and visual impairments. These would clearly need to be addressed prior to implementation in real-world clinical practice. For example, approved processes and safeguards would be in place to protect anonymity. Also, clinical staff raised concerns that some patients may be too weak to squeeze a hand, potentially limiting feasibility and validity of this response method [25]. Of note, the SAVEAHAART, a key part of the CAM-ICU for assessing inattention which relies on hand squeezing as method of response, has not been validated on its own for use in ICU patients. The suitability of this response method in ICU patients needs to be further explored. Nonetheless, overall the DelApp-ICU was found to be a promising tool for detecting and monitoring arousal and attention impairments in delirium in the ICU. The use of DelApp and variant tests in several completed and ongoing research studies provides evidence supporting its feasibility [16,18,30-33].

A case-control design using a relatively small patient sample was employed and patients with a RASS score < -2 were excluded from the study, thereby limiting generalisability. The study was not sufficiently powered for the longitudinal analysis. We did not compare our findings to a validated measure of delirium severity. The CAM-ICU was used as a reference standard for delirium diagnosis. Some studies have reported moderate sensitivity of this tool [15]; hence milder forms of delirium may have been missed in the present study. Also, the researchers were aware of the CAM-ICU status of patients at the time of administering the DelApp-ICU assessment. Although the DelApp-ICU scoring does not rely strongly on subjective judgment, we cannot exclude that knowledge of delirium status may have influenced DelApp-ICU assessments to some extent. The DelApp-ICU requires active task participation, which is a limitation common to most delirium assessment methods. However, the DelApp-ICU offers more flexibility through selection of a suitable method of response and allows for scoring of markedly drowsy or agitated patients.

4.2. Implications for research and clinical practice

Being an initial exploratory study, further evaluation of the real-world reliability and acceptability of the DelApp-ICU as a diagnostic tool for detecting and monitoring attention impairments in delirium in the ICU is required. Additional qualitative staff interviews specifically focusing on the presentation and usability of the DelApp-ICU should be carried out. To assess diagnostic performance and inform optimal cut-points, a larger case-control study in ICU patients (including patients with a RASS score < —2) followed by a definitive study in a representative sample of ICU patients is required. Such studies should include a more comprehensive reference standard evaluation including detailed assessment of arousal, attention (from bedside assessment) and timing of onset (from informants, if available) [26], and should have

independent assessors administering the index test (DelApp-ICU) and reference standard assessment. Ultimately the DelApp-ICU assessment may have applications in both clinical practice and research, where it could supplement existing instruments or provide an alternative, particularly in cases where monitoring delirium severity over time is required. The scalar aspect of the DelApp-ICU may also be useful for determining who is at higher risk of delirium as poor premorbid attention has been associated with subsequent delirium [27].

The present study also provides preliminary support for the utility of the OSLA as a tool for detailed measurement of level of arousal in ICU patients, though the OSLA itself has yet to be formally validated in this population.

Our study fits with the growing trend towards using smartphone-based assessments for assessing cognitive function in patients [28]. A meta-analysis of healthcare-related apps indicated that the majority of applications were aimed at healthcare professionals focusing on disease diagnosis, and highlighted the use of smartphones as emerging technology for point-of-care testing in clinical practice [29]. Tools such as the DelApp-ICU may encourage and facilitate greater adoption of delirium detection and monitoring, which is a serious and yet under-researched problem affecting large numbers of ICU patients.

5. Conclusion

This is the first study to provide evidence for the feasibility and initial validity of a brief and simple smartphone-based test for quantifying impairments in level of arousal and attention associated with delirium in the ICU. The DelApp-ICU shows promise as an objective tool to assist detection of delirium in research, with potential applications in clinical practice after further development and formal validation.

Conflict of interest

AMJM holds patents on computerized devices and tests for measuring attention in delirium. For the remaining authors none were declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcrc.2018.08.019.

References

[1] Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, 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(21):2703–10.

[2] Ely EW, Shintani A, Truman B, Speroff T, Gordon SM, Harrell Jr FE, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA 2004;291(14):1753–62.

[3] Jackson JC, Gordon SM, Hart RP, Hopkins RO, Ely EW. The association between delirium and cognitive decline: a review of the empirical literature. Neuropsychol Rev 2004;14(2):87–98.

[4] Page V, Gough K. Management of delirium in the intensive care unit. Br J Hosp Med 2010;71(7):372–6.

[5] Pisani MA, Kong SY, Kasl SV, Murphy TE, Araujo KL, Van Ness PH. Days of delirium are associated with 1-year mortality in an older intensive care unit population. Am J Respir Crit Care Med 2009;180(11):1092–7.

[6] Barr J, Fraser GL, Puntillo K, Ely EW, Gelinas C, Dasta JF, et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the Intensive Care Unit: executive summary. Am J Health Syst Pharm 2013;70(1):53–8.

[7] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington DC: APA; 2013.

[8] Bledowski J, Trutia A. A review of pharmacologic management and prevention strategies for delirium in the intensive care unit. Psychosomatics 2012;53(3):203–11.

[9] Flaherty JH, Gonzales JP, Dong B. Antipsychotics in the treatment of delirium in older hospitalized adults: a systematic review. J Am Geriatr Soc 2011;59(Suppl. 2): S269–76.

[10] Hoyle G, Sharma V, MacLullich A, Seymour DG. Clinical aspects of delirium. J R Coll Physicians Edinb 2008;38(2):154–7.

[11] Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). Crit Care Med 2001;29(7):1370–9.

[12] Bergeron N, Dubois MJ, Dumont M, Dial S, Skrobik Y. Intensive care delirium screening checklist: evaluation of a new screening tool. Intensive Care Med 2001;27(5): 859–64.

[13] Luetz A, Heymann A, Radtke FM, Chenitir C, Neuhaus U, Nachtigall I, et al. Different assessment tools for intensive care unit delirium: which score to use? Crit Care Med 2010;38(2):409–18.

[14] McNicoll L, Pisani MA, Ely EW, Gifford D, Inouye SK. Detection of delirium in the intensive care unit: comparison of confusion assessment method for the intensive care unit with confusion assessment method ratings. J Am Geriatr Soc 2005;53(3): 495–500.

[15] van Eijk MM, van den Boogaard M, van Marum RJ, Benner P, Eikelenboom P, Honing ML, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. Am J Respir Crit Care Med 2011;184(3):340–4.

[16] 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):1244–53.

[17] 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): 1334–40.

[18] Green C, Hendry K, Wilson ES, Walsh T, Allerhand M, MacLullich AMJ, et al. A novel computerized test for detecting and monitoring visual attentional deficits and delirium in the ICU. Crit Care Med 2017;45(7):1224–31.

[19] Tieges Z, Stiobhairt A, Scott K, Suchorab K, Weir A, Parks S, et al. Development of a smartphone application for the objective detection of attentional deficits in delirium. Int Psychogeriatr 2015;27(8):1251–62.

[20] Sessler CN, Gosnell MS, Grap MJ, Brophy GM, O'Neal PV, Keane KA, 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):1338–44.

[21] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13(10):818–29.

[22] R Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.

[23] Kean J, Ryan K. Delirium detection in clinical practice and research: critique of current tools and suggestions for future development. J Psychosom Res 2008;65(3): 255–9.

[24] Boettger S, Nunez DG, Meyer R, Richter A, Schubert M, Jenewein J. Subsyndromal delirium in the intensive care setting: Phenomenological characteristics and discrimination of subsyndromal delirium versus no and full-syndromal delirium. Palliat Support Care 2017:1–11.

[25] Oxenboll-Collet M, Egerod I, Christensen V, Jensen J, Thomsen T. Nurses' and physicians' perceptions of confusion Assessment Method for the intensive care unit for delirium detection: focus group study. Nurs Crit Care 2016;23(1): 16–22.

[26] Tieges Z, Evans JJ, Neufeld KJ, MacLullich AM. The neuropsychology of delirium: advancing the science of delirium assessment. Int J Geriatr Psychiatry 2017. https://doi.org/10.1002/gps.4711.

[27] Fann JR, Roth-Roemer S, Burington BE, Katon WJ, Syrjala KL. Delirium in patients undergoing hematopoietic stem cell transplantation. Cancer 2002;95(9): 1971–81.

[28] Brouillette RM, Foil H, Fontenot S, Correro A, Allen R, Martin CK, et al. Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. PLoS One 2013;8(6):e65925.

[29] Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012;12:67.

[30] Neerland BE, Hov KR, Bruun Wyller V, Qvigstad E, Skovlund E, AM MacLullich, et al. The protocol of the Oslo Study of Clonidine in Elderly Patients with Delirium; LUCID: a randomised placebo-controlled trial. BMC Geriatr 2015;10:15–7.

[31] Shenkin SD, Fox C, Godfrey M, Siddiqi N, Goodacre S, Young J, et al. Protocol for validation of the 4AT, a rapid screening tool for. delirium: a multicentre prospective diagnostic test accuracy study. BMJ Open 2018;10(2):e015572.

[32] Davis D, Richardson S, Hornby J, Bowden H, Hoffmann K, Weston-Clarke M, et al. The delirium and population health informatics cohort study protocol: Ascertaining the determinants and outcomes from delirium in a whole population. BMC Geriatr 2018;18(1):45.

[33] Yoshimura, A. et al. Altered cortical brain activity in end stage liver disease assessed by multi-channel near-infrared spectroscopy: Associations with delirium. Sci Rep. 7 (1):9258. doi: https://doi.org/10.1038/s41598-017-10024-7