Development of Android Apps for Cognitive Assessment of Dementia and Delirium

Alexander J. Weir[1]-*IEEE Member*, Craig A. Paterson[2], Zoe Tieges[3], Alasdair M. MacLullich[3], Mario Parra-Rodriguez[4], Sergio Della Sala[4] and Robert H. Logie[4]

***Abstract*— The next generation of medical technology applications for hand-held portable platforms will provide a core change in performance and sophistication, transforming the way health care professionals interact with patients. This advance is particularly apparent in the delivery of cognitive patient assessments, where smartphones and tablet computers are being used to assess complex neurological conditions to provide objective, accurate and reproducible test results. This paper reports on two such applications (apps) that have been developed to assist healthcare professionals with the detection and diagnosis of dementia and delirium.**

I. INTRODUCTION

The ability to obtain an accurate and reproducible cognitive assessment of patients remains a significant challenge for healthcare professionals. This is particularly the case when assessing the elderly or infirm for the presence of dementia or delirium (DL).

In the case of dementia, it has been demonstrated that, in patients with Alzheimer's disease (AD), Dual-Task (DT) performance (the ability to perform two independent tasks simultaneously) is impaired relative to normal controls as wells as in patients with mild cognitive impairment (MCI) [1]. This deficit is known to become more pronounced with increasing severity of the disease [2] [3]. Furthermore, it has also been demonstrated that Working Memory Binding (WMB) (the temporary retention in working or short-term memory of the shape and/or colour of complex objects [4]) is also sensitive to AD, and may be a pre-clinical marker for familial AD (FAD) [5].

DL is an acute, severe deterioration in mental functioning with severe consequences for patient outcome. It is known to affect at least 1 in 8 of general acute hospital inpatients, and at least 1 in 3 intensive care unit (ICU) inpatients. It is associated with considerable patient and carer distress and leads to longer stays in hospital. Moreover, due to similarities between DL and the symptoms of dementia, it is important to screen for DL when assessing other cognitive conditions such as AD, particularly as there are fundamental differences in the associated care pathways of these conditions. Despite its importance, DL is grossly under-detected in acute hospitals, with less than 1 in 3 diagnosed. The main cause of these low detection rates is the lack of a validated, reproducible, objective test. However, it has been shown that one of the core features of DL is inattention [6] and an abnormal level of arousal (LoA), conditions for which a validated technique for detection has been developed [7].

The aim is to develop a suite of cognitive assessment apps for use on smartphones or tablet computers. This paper describes the initial development of this suite; two cognitive assessment apps which implement test methods for DT and WMB in the case of AD, and abnormal LoA and inattention in the case of DL. The test method, analysis and validation process of each app is described firstly for the Edinburgh Dementia App (EDA) in section II, and secondly the Delirium App (DELAPP) in section III. Finally, section IV provides a brief conclusion.

II. EDINBURGH DEMENTIA APP (EDA)

The EDA supports two paradigms that are clear candidates as clinical markers for discrimination between AD and MCI. These are the classical DT approach and the WMB approach [8]. Each is selectable from the main menu of the EDA. The app implements each on a 10-inch Android tablet computer running the Google Android OS version 4.0 or above. The app supports encryption and Personal Identification Number (PIN) protection to ensure security of the Structured Query Language (SQL) patient results database. In each test case the tablet is held at a comfortable reading distance from the test participant, and the use case includes a wireless Bluetooth keyboard for operator response entry and a stylus for use by the participant during the target tracking tasks.

A. Method

1) Dual-Task (DT) Paradigm: In the case of the DT test, the participant's baseline performance is assessed in terms of digit span and target tracking span.

For digit span, the participant is presented initially with a list of three spoken digits, delivered in English by a male voice using the Android OS text-to-speech (TTS) engine at a delivery rate of one digit per second. The participant is instructed to verbally recall the list. There is no time limit for the recall. Each recalled digit is entered on the keyboard

[1]A.J. Weir is with the Medical Devices Unit, Department of Clinical Physics, Southern General Hospital, Glasgow, United Kingdom and the School of Engineering and Physical Science, Heriot-Watt University, Edinburgh, United Kingdom. alexander.weir@nhs.net

[2]C.A. Paterson is with the Medical Devices Unit, Department of Clinical Physics, Southern General Hospital, Glasgow, United Kingdom. craigpaterson@nhs.net

[3]Z. Tieges and A.M. MacLullich are with the Geriatric Medicine Unit, University of Edinburgh, Edinburgh, United Kingdom. zoe.tieges@ed.ac.uk and a.maclullich@ed.ac.uk

[4]M. Parra-Rodriguez, S. Della Sala and R.H. Logie are with the Department of Human Cognitive Neuroscience, University of Edinburgh, 7 George Square, Edinburgh, United Kingdom. mprodri1@staffmail.ed.ac.uk, sergio@ed.ac.uk and rlogie@staffmail.ed.ac.uk

978-1-4244-7929-0/14/$26.00 ©2014 Crown

by the operator. If three out of four digit sequences are correctly recalled, the sequence length is increased by one digit and the process begins again. This process continues until the participant is unable to correctly recall at least three sequences. The participant's digit span is recorded as the maximum sequence length at which a participant is able to correctly remember three out of four sequences.

For tracking span, the participant is presented with a blue bug on a black tablet screen, as shown in figure 1a. The blue bug moves around the screen, changing directions randomly. The participant is asked to track the bug using a stylus. Whenever the stylus is on the bug, the bug turns red as shown in figure 1b. Whenever the stylus moves off the bug the colour returns to blue. The bug moves slowly at first, gradually increasing in speed in $20s$ intervals until the participant is unable to track the bug for more than 60% of the time. The participant's tracking span is recorded as the maximum speed at which a participant is able to track the bug for more than 60% of the time.

Once the digit span and tracking span of a participant are established using the technique described, they are confirmed by re-testing the participant with a single digit span test followed by a single tracking span test.

Finally, participants complete the DT test by performing the digit task together with the tracking task simultaneously for $90s$ at their measured digit and tracking spans.

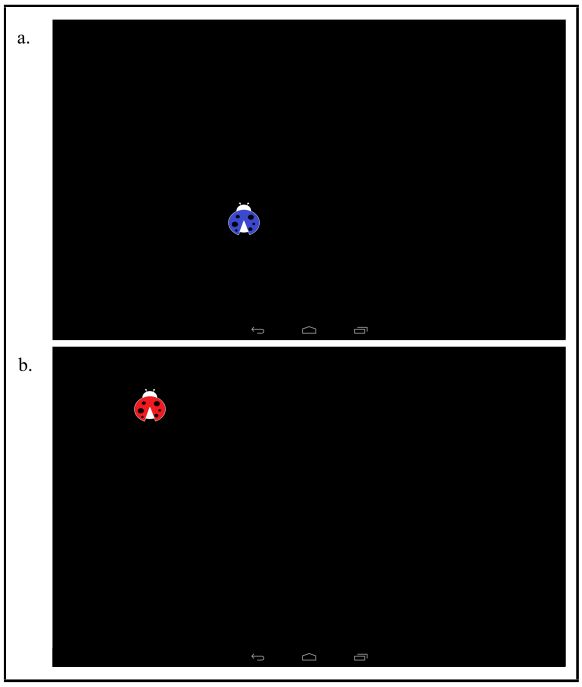

Fig. 1. The adaptive tracking test: The target bug image moves in random directions at a slowly increasing speed around the screen. (a) When the test participant has their finger/stylus off the target, the bug turns blue. (b) When the test participant has their finger/stylus on the target, the bug turns red.

2) Working Memory Binding (WMB) Paradigm: The WMB test is divided into two subtasks; a perceptual binding task and a WMB task. The perceptual binding task provides a screening function for the WMB task.

a) Perceptual Binding task: In this task the participant uses a visual search to interpret two arrays of two polygons presented simultaneously on the tablet screen, as shown in figure 2a. The screen is divided symmetrically by a black horizontal line. In 50% of the trials, the arrays consist of polygons of the same colour presented at different locations. In the remaining 50% of trials, items in one array will swap colours. The participant, answering as quickly and accurately as they can, is asked whether the two arrays are the "same" or "different". The arrays remain on the screen until the participant responds. Once a response is given, the operator touches the screen and the view shown in figure 2b is presented. The operator presses the corresponding response button to record the answer. A score above 80% accuracy is required to proceed to the WMB task. A score below 80% requires further visual testing (e.g., colour vision, visual acuity, etc.) to decide whether further memory binding testing is meaningful.

b) Working Memory Binding Task: This task assesses the binding of colour and shape. As before, each trial consists of two sets of coloured polygons; a study display is shown briefly on the screen. This is followed by a test display. The test display is identical to the study array in terms of shape and colour 50% of the time, although location may vary. Location variance ensures that the position of each item on screen cannot be used to help remember the shape or shape-colour combinations. Each trial consists of the following sequence; a central fixation ($0.5s$), followed by a study display ($2s$), followed by a blank display ($1s$), followed by the test display. The participant, answering as quickly and accurately as they can, is asked whether the study display and the test display were the "same" or "different". Once a response is given, the operator touches the screen and then presses the corresponding response button to record the answer. There is then a $1s$ delay before the next study display appears. The WBM task consists of 64 trials in total; 32 trials with black polygon combinations (testing memory binding of shape only, as shown in the trial sequence depicted in figure 3a), followed by 32 trials with coloured polygon combinations (testing memory binding of shape and colour, as shown in the trial sequence depicted in figure 3b).

B. Results

The EDA software has been fully tested in terms of function and reliability using a JUnit test framework within the Eclipse integrated development environment (IDE). It has completed usability trials to assess the suitability of the app as a means of early screening for dementia within primary care, as part of a standard GP consultation within a medical centre or care home, or during a hospital outpatient assessment. The underlying test cases are based on the synthesis of several significant studies over the past two-decades [9]–[12], [14], [15]. From these studies it is known that DT provides an objective measure, does not show practice based improvements, is unaffected by age and it automatically adjusts to an individual's level of ability. Yet, it provides a specific measure of the deficits associated with AD. However,

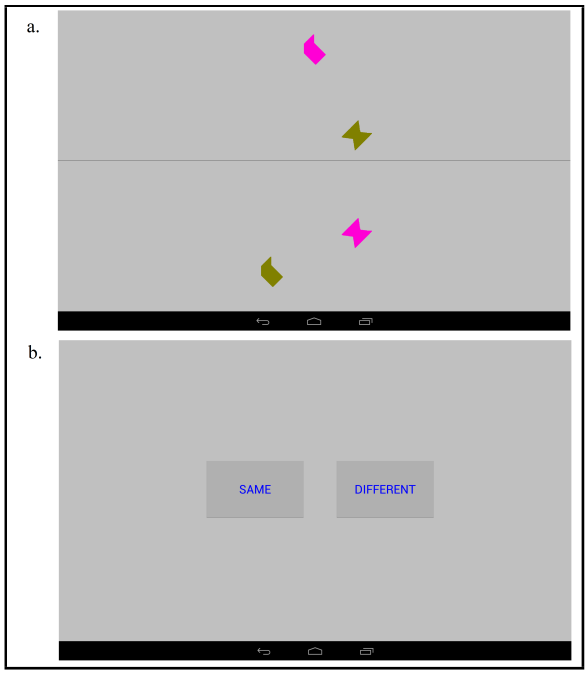

Fig. 2. The WMB test: (a) The test participant is presented with a screen divided horizontally by a black line. Each half shows an array of two coloured polygons. The participant judges whether the polygon-colour combinations in both arrays are the same or different. (b) The operator touches the screen and enters the participant's response.

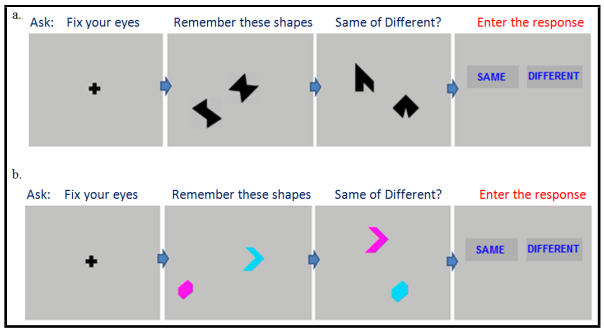

Fig. 3. (a) An example trial sequence for the *memory of shape test*. (b) An example trial sequence for the *memory of shape and colour test*.

it has been shown that the sensitivity of the DT test may not be ideal [13]. The WMB test was developed to address the sensitivity problems of the DT test. With the added advantage that it is language independent, the WMB test has been proven to clearly differentiate the deficits associated with AD. Furthermore, it has shown much promise by observing selective deficits in WMB of asymptotic carriers of the E280A presenilin 1 AD gene mutation, the cause of the most common form of dominant early onset FAD [5]. One disadvantage of the WMB approach is the requirement for normal visual function; participants must be able to clearly distinguish both shape and colour to perform an assessment. A suitable visual acuity and colour blindness test must be used to assess the participant's suitability for assessment.

III. DELIRIUM APP (DELAPP)

Although several validated tools for the assessment of DL are available, there is a general lack of detailed research on methods of objectively measuring inattention. Of those tests that are available, most do not adequately discriminate DL from dementia. Furthermore, the lack of a reliable measure of grading the severity of inattention and monitoring attentional function over time has contributed to the under-detection of DL in clinical practice. The DELAPP aims to address these issues by providing an objective measure of LoA and inattention. Based on a methodology developed using the Edinburgh Delirium Test Box MkII (EDTB2) [7] (a neuropsychological test unit for the objective measurement of inattention in DL), the app is designed for a 5-inch Android smartphone running the Google Android OS version 2.3 and above. The smartphone is held with the display positioned at a comfortable viewing distance from the test participant. The participant is asked to observe stimulus on the smartphone screen and respond to verbal cues from the operator.

A. Method

The DELAPP provides a suite of simple cognitive tests, including a visual acuity test, word building task and counting task. However, only the visual acuity and counting task are discussed in this study. The visual acuity test consists of a small number of test cases to confirm the participant has no visual defects that might otherwise compromise their ability to complete the cognitive tests. The visual acuity test must be performed as a pre-test before proceeding with the counting task.

The operator begins the counting task by assessing the alertness of the patient. An alertness score is compiled based on whether the participant can keep their eyes open for $10s$, say their name and visually track an object (e.g. a badge, phone etc.) for $5s$. The counting task consists of 8 trials, the first being a practice trial. The participant counts a sequence of stimuli (flashing white circles) presented on the screen. Each stimuli is presented for a duration of $1s$ and the time delay between stimulus varies based on a pre-defined design. As the trials progress, distracting shapes are included in the sequence around the stimuli and the time delay between flashes increase. The shape, size and frequency of the distractors is informed by previous studies [7]. Figure 4 shows views of the counting task.

B. Results

The DELAPP software has been fully tested in terms of function and reliability using a JUnit test framework within the Eclipse IDE. It has also completed usability and proof-of-concept trials to assess the suitability of the app for use in both general hospital wards and ICUs. Further testing has focused on the counting task in preference to other tasks available in the app. An initial feasibility study which compared the performance of the DELAPP counting task with the EDTB2 in 20 hospital patients demonstrated highly comparable performance. Further studies of the DELAPP in 156 general ward patients has shown that patients with

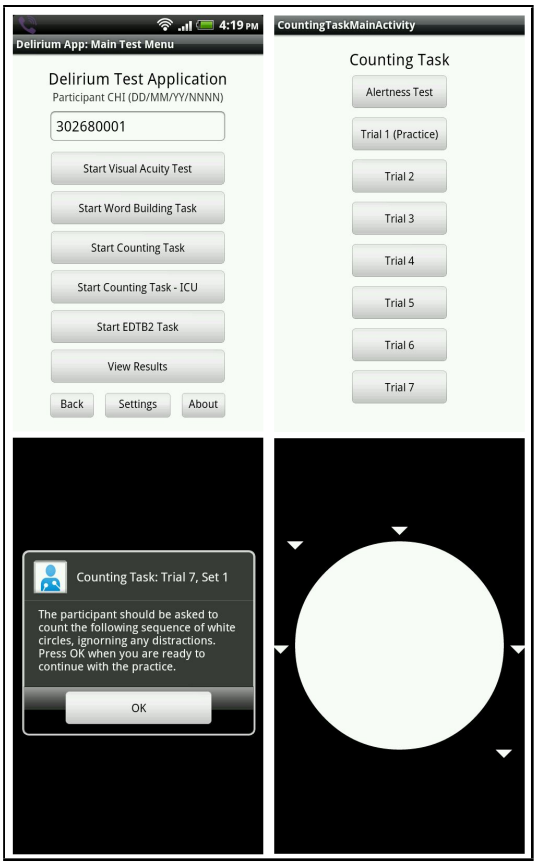

Fig. 4. The DELAPP counting task: The operator selects and begins a counting task, preparing the participant for the test. The participant is asked to count the number of appearances (flashes) of a central stimulus which appear at pre-designed time intervals surrounded by flickering distractors as the trials progress.

DL have statistically significantly lower scores compared to patients with dementia or cognitively normal volunteers. Using the standard binary classification system for statistical analysis of DL diagnosis and taking the Area Under the Receiver Operating Characteristics (AUC) as a measure of discrimination, the DELAPP achieved a test accuracy of 0.99, indicating excellent performance. A further feasibility study of the DELAPP in 47 patients in an ICU (20 diagnosed with DL) found that the DELAPP measure of discrimination achieved a test accuracy of 0.98, again indicating excellent performance. Additionally, it was observed that the DELAPP demonstrated good sensitivity to change in attentional functioning upon repeat testing in a subset of patients.

IV. CONCLUSION

Smartphones and tablet computers have been widely adopted in practice by hospital doctors, with estimates suggesting that over 80% of healthcare professionals use a smartphone in their daily professional capacity. The combination of flexibility, processing power and ease of use provides a powerful platform for the support of sophisticated medical technology applications. The EDA and DELAPP applications are excellent examples of this growing trend.

The EDA provides tests supporting both DT and WMB paradigms. These have been shown to provide objective measures of AD, whilst discriminating between the cognitive defects associated with AD and those of healthy ageing and other forms of dementia. Both methods show strong tolerance to 'practice' or learning effects, and are unaffected by the age or ability of the participant.

The establishment of a reliable and objective system of detection of DL is a major clinical priority. The results reported in this study suggest that the DELAPP is capable of detecting and quantifying LoA and inattention, and may provide a simple method of diagnosis for use in routine clinical care. The DELAPP is likely to have many advantages over existing test methods, including objectivity, automated scoring and ease-of-use. The DELAPP has been shown to have excellent test accuracy and good sensitivity as a method of detection of DL, both in general hospital wards and ICU. Furthermore, it provides the ability to discriminate DL from dementia and is capable of monitoring progression of the condition.

REFERENCES

[1] J.A. Foley, R. Kaschel, R.H. Logie, S. Della Sala. "Dual task performance in Alzheimer's disease, Mild Cognitive Impairment and normal ageing." *Neuropsychology*, vol. 26, pp. 340-348, 2011.

[2] A.D. Baddeley. *Working memory, thought and action*. Oxford: Oxford University Press, 2007.

[3] A.D. Baddeley,H.A. Baddeley,K.S. Bucks,G.K. Wilcock. "Attentional control in Alzheimers disease." *Brain*, vol. 124, pp. 14921508, 2001.

[4] M.A. Parra, S. Della Sala, R.H. Logie, A. Morcom. "Neural correlates of shape-color binding in visual working memory." *Neuropsychologia*, vol. 52C, pp. 27-36, 2014.

[5] M.A. Parra, S. Abrahams, R.H. Logie, L.G. Mendez, F. Lopera, S. Della Sala. "Visual short-term memory binding deficits in Familial Alzheimers Disease." *Brain*, vol. 133, pp. 2702-2713, 2010.

[6] L.J. Brown, C. Fordyce, H. Zaghdani, J.M. Starr, A.M. MacLullich. "Detecting deficits of sustained visual attention in delirium." *J Neurol Neurosurg Psychiatry*, vol. 82(12), pp. 1334-40, 2011.

[7] Z. Tieges, A. McGrath, R.J. Hall, A.M. Maclullich. "Abnormal level of arousal as a predictor of delirium and inattention: an exploratory study." *Am J Geriatr Psychiatry*, vol. 21(12), pp. 1244-53, 2013.

[8] S. Della Sala, J.A. Foley, M.A. Parra, R.H. Logie. "Dual tasking and memory binding in Alzheimers." *Journal of Alzheimer's Disease*, vol. S23, pp. 22-24, 2011.

[9] S. Della Sala, G. Cocchini, R.H. Logie, S.E. MacPherson. "Dual task during encoding, maintenance and retrieval in Alzheimer's disease and healthy ageing." *Journal of Alzheimer's Disease*, vol. 19, pp. 503-515, 2010.

[10] J.A. Foley, R. Kaschel, R.H. Logie, S. Della Sala. "Dual task performance in Alzheimer's disease, Mild Cognitive Impairment and normal ageing." *Archives of Clinical Neuropsychology*, vol. 26, pp. 340-348, 2011.

[11] R. Kaschel, R.H. Logie, M. Kazén, S. Della Sala. "Alzheimer's Disease, but not ageing or chronic depression, affects dual-tasking." *Journal of Neurology*, vol. 256, pp. 1860-1868, 2009.

[12] R.H. Logie, G. Cocchini, S. Della Sala, A.D. Baddeley. "Is there a specific executive capacity for dual task co-ordination? Evidence from Alzheimers Disease." *Neuropsychology*, vol. 18, pp. 504-513, 2004.

[13] S. Della Sala, J.A. Foley, N. Beschin, M. Allerhand, R.H. Logie. "Assessing Dual-Task Performance Using a Paper-and-Pencil Test: Normative Data." *Archives of Clinical Neuropsychology*, vol. 25, pp. 410-419, 2010.

[14] M.A. Parra, S. Abrahams, R.H. Logie, S. Della Sala. "Visual short-term memory binding in Alzheimer's disease and depression." *Journal of Neurology*, vol. 257, pp. 1160-1169, 2010.

[15] M.A. Parra, K. Fabi, R.H. Logie, S. Luzzi, S. Della Sala. "Short term memory binding deficits in Alzheimers Disease." *Brain*, vol. 132, pp. 1057-1066, 2009.