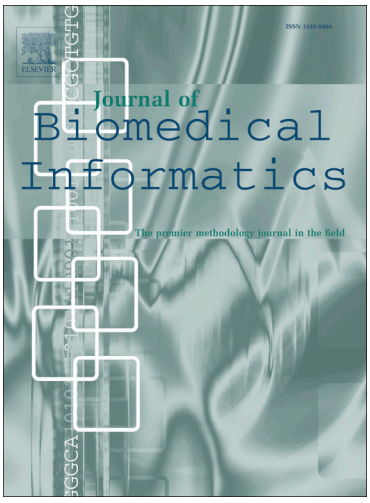
Journal Pre-proofs



Dementia Medical Screening using Mobile Applications: A Systematic Review with A New Mapping Model

Fadi Thabtah, David Peebles, Jenny Retzler, Chanchala Hathurusingha

PII: S1532-0464(20)30202-1
DOI: https://doi.org/10.1016/j.jbi.2020.103573
Reference: YJBIN 103573

To appear in: *Journal of Biomedical Informatics*

Received Date: 21 March 2020
Revised Date: 13 September 2020
Accepted Date: 14 September 2020

Please cite this article as: Thabtah, F., Peebles, D., Retzler, J., Hathurusingha, C., Dementia Medical Screening using Mobile Applications: A Systematic Review with A New Mapping Model, *Journal of Biomedical Informatics* (2020), doi: https://doi.org/10.1016/j.jbi.2020.103573

Dementia Medical Screening using Mobile Applications: A Systematic Review with A New Mapping Model

Fadi Thabtah
Manukau Institute of Technology
Auckland, New Zealand
Fadi.fayez@manukau.ac.nz
Corresponding Author

David Peebles
Department of Psychology
University of Huddersfield
Huddersfield, UK
d.peebles@hud.ac.uk

Jenny Retzler
Department of Psychology
University of Huddersfield
Huddersfield, UK
J.Retzler@hud.ac.uk

Chanchala Hathurusingha
Digital Technologies, Manukau Institute of Technology
Manukau, Auckland, New Zealand
Chanchala. Hathurusingha@manukaumail.com

Corresponding Author Details
Fadi Thabtah
School of Digital Technologies, Manukau Institute of Technology
Private Bag 94006, Manukau, Auckland, 2241, New Zealand.
E-mail: fadi.fayez@manukau.ac.nz.

Dementia Medical Screening using Mobile Applications: A Systematic Review with A New Mapping Model

Abstract

Early detection is the key to successfully tackling dementia, a neurocognitive condition common among the elderly. Therefore, screening using technological platforms such as mobile applications (apps) may provide an important opportunity to speed up the diagnosis process and improve accessibility. Due to the lack of research into dementia diagnosis and screening tools based on mobile apps, this systematic review aims to identify the available mobile-based dementia and mild cognitive impairment (MCI) apps using specific inclusion and exclusion criteria. More importantly, we critically analyse these tools in terms of their comprehensiveness, validity, performance, and the use of artificial intelligence (AI) techniques. The research findings suggest diagnosticians in a clinical setting use dementia screening apps such as ALZ and CognitiveExams since they cover most of the domains for the diagnosis of neurocognitive disorders. Further, apps such as Cognity and ACE-Mobile have great potential as they use machine learning (ML) and AI techniques, thus improving the accuracy of the outcome and the efficiency of the screening process. Lastly, there was overlapping among the dementia screening apps in terms of activities and questions they contain therefore mapping these apps to the designated cognitive domains is a challenging task, which has been done in this research.

Keywords: cognitive mapping, dementia, MCI, screening methods, mobile apps, systematic review, machine learning, neurodegenerative areas

1. Introduction

Dementia is associated with symptoms that include a progressive worsening of memory and learning ability as well as decline in other cognitive areas (Ernst & Hay, 1999). Dementia also includes other mild cognitive impairments s(MCIs), Alzheimer's disease (AD), Parkinson's disease, and Huntington's (Gustafson, 1996). In 2019, 5.6 million Americans aged above 65 years, and 0.2 million aged below 65 years were diagnosed with dementia conditions. These numbers are expected to increase in the coming years at a growth rate of 32% for people aged above 85 (Alzheimer's Association, 2019a). In America in 2019, 121,404 deaths resulted directly from AD or other types of dementia (Alzheimer's Association, 2019b). Thus, identifying dementia traits at the earliest stage through screening is vital to initiate the intervention process, manage the disease effectively, and grant individuals and their family members access to relevant healthcare services (Moon, Badana, Hwang, Sears, & Haley, 2019; Brodaty, Low, Gibson & Burns, 2006).

Dementia screening can be defined as the process by which an individual, who might be in the prodromal stage of dementia, is verified as a case of dementia using neuropsychological tests (Panegyres, Berry, & Burchell, 2016). There

are many dementia screening methods that can be administered by individuals, caregivers, or healthcare professionals. These provide a quick evaluation of cognitive and functional aspects such as memory, attention, concentration, orientation, visuoconstructional ability, language, and calculation. Depending on the scores obtained, these methods can pinpoint dementia traits and possible impairments in the individual undergoing the test. The Abbreviated Mental Test (AMT; Hodkinson, 1972), the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), and Addenbrooke's Cognitive Examination- Revised (ACE-R; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006) are examples of such dementia screening methods.

There are several mobile apps available to assist individuals already diagnosed with cognitive dysfunctions like dementia. Most of them focus on assisting patients with brain training, improving their memory power, and preventing negative incidents associated with cognitive impairment. However, few mobile apps have been developed to assist individuals and medical professionals in screening these conditions. Some dementia screening methods are now available on mobile platforms making them more accessible to diagnosticians, patients, caregivers, researchers, and professional medical staff.

Although several studies have been carried out on traditional dementia and MCI screening tests, particularly regarding performance, very little research has focused on evaluating mobile-based dementia screening tests. For instance, a study by Brodaty et al., (2006) compared non-app-based screening medical assessments that general practitioners often use, based on numerical criteria such as sensitivity, specificity, and time. Another study by Pang and Kwong (2015) evaluated apps to describe those available in 2015 and possible areas for further development, but was restricted to mild-to-moderate dementia and had a primary focus on cognitive simulation therapy and managing virtual daily living activities. Berauk et al., (2017) assessed features related to mobile health applications for elderly people using qualitative parameters like Internet connection, information of disease, assessment tools, medical calculator, dosage recommendations, clinical updates, drugs interaction checker, and information on disease management.

Little research, therefore, has been systematically carried out to evaluate the currently available dementia screening apps with respect to inclusion criteria such as accessibility, fulfilment of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), and coverage of cognitive areas. The few available mobile-based review studies on dementia screening tools are not comprehensive, leaving an area which this research aims to address (Groppell et al., 2019; Berauk et al., 2017; Yamagata & Kowto, 2013). Therefore, in this research we evaluate and systematically review dementia screening mobile apps available on both Apple and Android platforms. The evaluation criteria are coverage of DSM-5 neurocognitive domains, validity, reported performance (sensitivity and specificity), and inclusion of ML & AI techniques. More importantly, we propose a mapping between screening apps' available test and the DSM-5 neurocognitive domains in order to help clinicians identify the apps that cover multiple cognitive domains which will be more appropriate for dementia screening. Section 3.1 gives further details.

This study is of benefit to all parties including the patients, caregivers, and clinicians as it will help them to ascertain the most suitable, best performing, and valid dementia screening apps. The contribution of the paper is two folds:

- Systematic evaluation of dementia screening tools based on mobile platforms using critical criteria in order to identify the ones that are valid, cover most neurocognitive domains and accurate
- Proposing a mapping model between the six cognitive domains defined by DSM-5 with the available dementia related cognitive tests available on the mobile platform thus assessing clinicians in adopting the appropriate dementia screening tool

The paper consists of five main sections: Section two identifies and reviews prevailing screening apps. Section three systematically analyses each screening tool using inclusion and exclusion criteria. Lastly, section four highlights future works and limitations and section 5 concludes the findings.

2. Literature Review

A comprehensive search was conducted on the Google and Android Stores and iTunes search engine to identify available apps related to dementia and MCI screening. The initial search was based on a few key words: dementia, dementia screening, dementia tests, Alzheimer's screening, MCI exams, cognitive tests, mental status assessment, dementia diagnosis tools, and neurocognitive tests. Figure 1 shows the process of selecting the apps based on the inclusion and exclusion criteria (Table 1). The initial search yielded 275 apps (210 Android and 65 Apple Store apps). After removing duplicates, the remaining apps were checked separately to determine whether they were utilized for MCI, AD, or dementia screening. Most of the evaluated apps were games and informative apps to assist the patient to improve skills or cognitive function and provide relevant information on dementia and other cognitive impairments. Therefore, out of 275 apps, 214 were excluded as they didn't concern screening and were hence outside of this research scope. Table 1 shows inclusion and exclusion criteria used to further narrow the search. Further, 41 apps further were discarded as they are not based on valid dementia medical test, and just 20 apps retained (details of these screening apps are given in Table 2 (Thabtah et al., 2020)).

Table 1: Exclusion and Inclusion Criteria Used

Inclusion Criteria for Apps	**Exclusion Criteria for Apps**
Focuses on screening	Focuses on simulation therapy
Focuses on dementia, AD, and MCI	Focuses on other cognitive disorders
Available in English language	Not available in English language
Based on a valid medical test	Not based on a valid medical test
	Duplicates found in Android and Apple Stores

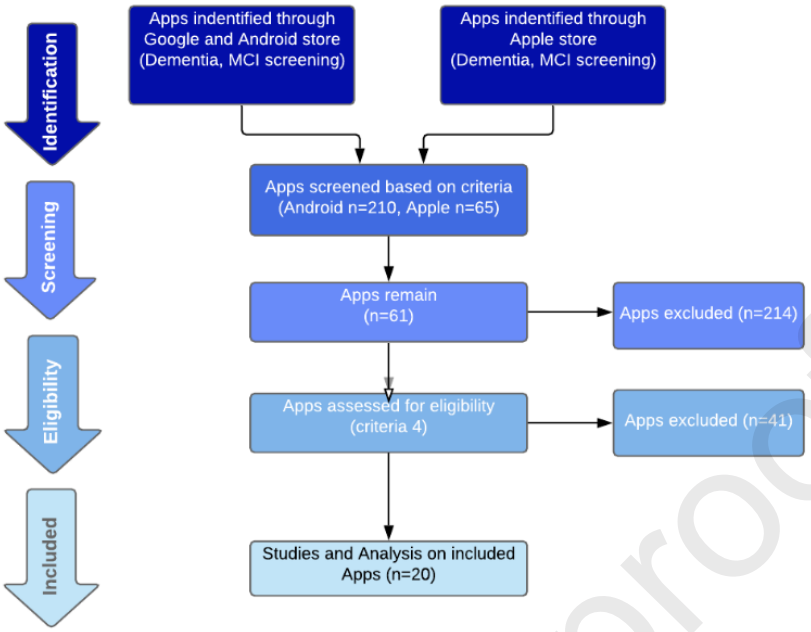


Fig. 1: The process of collecting the screening apps

2.1 Apps Based on a Single Medical Assessment Method

MOBI-COG (Nirjon, et al, 2014), a screening app based on the Mini-Cog Dementia Screening Test (Borson et al, 2000) aims to automatically identify dementia through three different tasks. This app can be administered by either the patient or the caregiver and takes three minutes to complete. It generates a score ranging from 1–10 and requires an expert to interpret the results. According to (Nirjon et al., 2014), the app is capable of screening dementia with 99.53% accuracy and has a rating of 8.5/10 on the Android App Store.

BrainTest® (BrainTest Inc, 2013) is based on the Self-Administered Gerocognitive Exam (SAGE) (Scharre, et al., 2010) and available in both Android and iOS operating systems. The app is designed to be administered by the patient or the caregiver and takes 10–15 minutes to complete. According to Scharre, Chang, Nagaraja, Vrettos, and Bornstein (2017), BrainTest® is capable of discriminating AD with a sensitivity and specificity of 71% and 90% respectively.

ACE (Hodges & Larner, 2017) is an app developed to be administered specially by healthcare professionals and is the automated version of Addenbrooke's Cognitive Examination-3 (ACE-3) medical examination (Noone, 2015). ACE supports iPads and evaluates the attention, memory, fluency, language, and visuospatial functioning capabilities of the individual then presents a score out of 100. According to Bruno and Vignaga (2019), ACE is capable of screening AD with a sensitivity of 93%–100% and specificity of 96%–100% and has a rating of 4+ on iTunes, despite some negative reviews regarding its functionality.

CAIDE-DRS (Sindi, et al., 2015) is an iOS app that uses the CAIDE risk score (cardiovascular risk factors, aging, and incidence of dementia; Kivipelto, et al., 2006) to evaluate an individual's cognition based on their biographic

information, systolic blood pressure (BP), body mass index, cholesterol level, and level of physical activity. CAIDE-DRS has a test that takes 5–10 minutes to complete and produces a score out of 15. Butcher (2007) revealed that CAIDE-DRS can determine the user's risk of developing AD within the next 20 years with sensitivity and specificity of 77% and 63% respectively. CAIDE-DRS gained a 4+ rating on iTunes. Similar to CAIDE-DRS, the DRT app (Google, 2019b) is also based on the CAIDE risk score (Kivipelto, et al., 2006) and is available in both Android and iOS formats making it more accessible than CAIDE-DRS. DRT has a rating of 4.8/5 on the Google Appstore with over 1000 recorded downloads.

MoCA (Apple Inc, 2019b; Nasreddine et al., 2019) is an iOS app based on the Montreal Cognitive Assessment which evaluates executive functions, language, attention, memory recall, concentration, and time and place awareness of the individual, to diagnose many brain and memory conditions including AD. The test takes approximately 10 minutes to complete and presents a score out of 30, where score <26 is interpreted as normal cognition. MoCA is used diagnose AD and other types of MCIs with a sensitivity and specificity of 81% and 86% respectively (Julayanont, et al., 2015). It has gained 4+ a rating on App Store.

eSLUMS (Chewy Logic LLC, 2019) is the digital version of the SLUMS medical test and designed to be administered by a healthcare professional. The app evaluates the user's brain health within 7–10 minutes and generates a score out of 30 where score >27 indicates normal cognition and score <20 indicates severe impairment. A study by Kansagara and Freeman (2010) revealed that SLUMS identifies AD traits with sensitivity and specificity ranging between 98–100%.

Two Android apps: **MMSE** (Google, 2016), and the Dementia & Alzheimer's Memory Diagnosis Test (**DAMDT;** Google, 2017), have been developed based on the MMSE medical test. Both apps generate a score out of 30 where a score >25 indicates normal cognition and score <25 indicates limitations on certain cognitive areas. Both the apps have a 3+ rating on the App Store, and DAMDT has over 1000 recorded downloads.

The **6CIT** app is based on the Six Cognitive Item Test medical assessment (Callahan et al, 2002). This test takes five minutes to complete and produces a score out of 28 where score <7 indicates normal cognition and a score >8 indicates an abnormality in cognitive processing and behavioural changes of the individual. The 6CIT app has a sensitivity range from 78–90% and specificity of 100% (Callahan et al., 2002).

2.2 Apps Based on Multiple Medical Assessment Methods

DementiaScreener (Mundt, Freed, & Greist, 2000) is an Android app based on the Symptoms of Dementia Screener (SDS; Flaherty, Midden, & Mast, 2019) and AD8 Dementia Screening Interview (Galvin, et al., 2005) medical exams taking five minutes to complete. SDS generates a score based on the repetitive behaviours, memory capacity, emotions, attention, and problem-solving skills of the individual. A score <4 indicates low risk and a score >4 indicates a high risk of developing AD. AD8 evaluates the behavioural changes of the individual where a score of 0–1 indicates

normal cognition, and a score >2 indicates impairment. The app gained a 2.7 rating on App Store with over 1000 recorded downloads.

CDD (Sangha, George, Winthrop, & Pancha, 2015) is a UK-NHS clinically accepted iOS app involving multiple medical assessment methods including the Confusion Assessment Method (CAM; Inouye, et al., 1990), Six Cognitive Item Screener (Callahan et al., 2002), the AMT, MoCA, and MMSE to screen several MCIs and dementia conditions. The CDD app is specially designed to be practised on orthopaedic trauma patients and can be used in clinical setup by clinicians. It has 4+ rating on the App Store.

BrainCheck (Ehrensperger et al., 2014) is an iOS app that comprises three tasks: Answering three questions (Holsinger, Deveau, Boustani, & Williams, 2007), the Clock Drawing Test (CDT; Friedman, et al., 1994), and seven informant questions from the IQCODE medical method (Jorm, Scott, & Jacomb, 1989). The app takes 10 minutes to complete and has a sensitivity and specificity of 97.4% and 81.6% respectively (Ehrensperger et al., 2014).

ALZ (Berauk et al., 2017) is an iOS app that uses the: Mini-Cog Dementia Screening Test, Memory Impairment Screen (MIS; Buschke et al., 1999), CDT, General Practitioner Assessment of Cognition (GPCOG; Brodaty et al., 2002), MoCA, IQCODE, AD8 (Galvin et al., 2005); SLUMS, Functional Activities Questionnaire (FAQ; Tappen, Rosselli, & Engstrom, 2009), Geriatric Depression Scale (Yesavage et al., 1982), Hachinski Ischemic Score (HIS; Johnson et al., 2014), and the Katz Index of Independence in activities of daily living as a functional assessment medical test (Aske, 1990) to screen for AD. This app is suitable for thorough medical screening of dementia in clinical setup and can be administered by a clinician.

DementiaTest (Thabtah, 2019) is an app based on the 6-item Cognitive Impairment (6-CIT) and the Structured Clinical Interview (SCIDS) medical tests (Callahan et al., 2002; Ouimette & Klein, 1995) and available on both Android and iOS formats. It consists of one self-administered questionnaire which produces a score out of 26 and a caregiver administered questionnaire which produces a score out of 36. Both questionnaires interpret a lower score as indicating lower risk of developing AD, and a higher score as indicating a higher risk of developing AD. The app takes 5–10 minutes to complete and has a rating of 3+ on the App Stores.

Cognitive Exams (Google, 2019a) is an Android app that uses MMSE, CDT, the Geriatric Depression Scale (Yesavage et al., 1982), Katz Examination, Walk Test of six minutes (Abbott et al., 2004; AbilityLab, 2019) and Fluency Test (Caramelli et al., 2007) to evaluate various functional and cognitive capabilities of the patients. This app can only be administered by a medical professional, takes approximately 45 minutes to complete and has a rating of 4.4 on the App Store with over 10,000 recorded downloads.

2.3 Apps Based on Non-conventional Methods

DTRCT (Apple Inc, 2019a) is an iOS app that uses a novel self-administered questionnaire based on several neurological test algorithms and interactive tools to evaluate the memory, attention, and logical thinking skills of

the user. The test takes five minutes to complete and provides a few other disease management services other than screening (Comerford et al, 2002).

BCI (BrainCheck, 2019) is a U.S. Food and Drug Administration (FDA) registered class III medical app available on iTunes. The app uses five games: Flanker Task (Fan et al, 2003), the Digit Symbol Substitution Task (Monte, Geffen, May, & McFarland, 2010), the Stroop Task (Melara, & Algom, 2003), the Trail Making Test (TMT; Cicerone & Azulay, 2002), and balance and coordination (Greenwald et al., 2001), The Immediate and Delayed Recall Tests were developed based on gold-standard neurocognitive tests to evaluate cognitive processing, executive functions, visual attention, immediate recall and the delayed recall abilities of the user. This app takes 10-minutes and has a sensitivity and specificity of 83% and 87% respectively.

Cognity (Inoven, 2018) is the only app that uses AI technology to screen for AD through analysing a photo of a clock drawing done by the user in conjunction with the Mental Status Examination (MSE; Snyderman & Rovner, 2009) interactive tool. The test takes 10-15 minutes to complete and produces a score out of 30. It has pre-recorded videos to interpret the results and sensitivity and specificity ranges between 71–92% and 52–96% respectively with a recorded rating of 3+.

The **DST** (Google, 2019c) app is available on both Android and iOS formats and used in dementia screening to evaluate the executive functions, memory, verbal fluency, attention, and orientation skills of the user. The test takes 10–15 minutes to complete and generates a score out of 30. A score >29 indicates normal cognition whereas a score <28 indicates impairment in some cognitive areas. According to Google (2019c), this app has a sensitivity of 96% and 3+ rating on the App Store.

Table 2: Summary of all considered AD screening Apps (Thabtah et al., 2020)

App	Target	Rating	Downloads	Apple iTunes	Google play Store	Video	Image	No of Questions	Time	Free/ Paid	Cost	Language	Reference
MOBI-COG	Anyone	8.5/10	N/A	X	√	X	√	3	3 mins	Paid	N/A	English	Borson, Scanlan, Brush, Vitaliano & Dokmak, (2000)
BrainTest	Anyone	4.5/5	N/A	√	√	√	√	Unknown	10-15 mins	Paid	$39.99	English	Scharre, Chang, Nagaraja, Vrettos, & Bornstein, 2017
ACE	Health care professionals	4+	N/A	√	√	X	√	24	5-10 mins	Free	N/A	English	(Noone, 2015)
CAIDE-DRS	Age 40 to 65	4+	1	√	X	X	X	8	5-10 mins	Free	N/A	English, Finnish, French, German, Russian, Spanish, Swedish	(Kivipelto, et al., 2006)
DRT	Age 40+	4.8	1000+	√	√	X	X	8	2-3 mins	Free	N/A	English	(Kivipelto, et al., 2006)
Dementia Screener	Anyone	2.7	1000+	X	√	√	X	11	5 mins	Free	N/A	English	(Galvin , et al., 2005) (Flaherty, Midden, & Mast, 2019)

MoCA	Anyone	4+	1	√	X	X	√	11	10 mins	Piad	$10.00	English, Danish, Dutch, Finnish, French, German, Italian, Polish, Portuguese, Spanish, Swedish	Ziad S. Nasreddine, et al., 2005
CDD	Orthopaedic trauma Patients	4+	N/A	√	X	X	X	10	10 mins	paid	N/A	English	(Inouye, et al., 1990), (Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002), (Hodkinson H. , 1972), (Nasreddine Z. , 2019) (Folstcin, Folstein, & McHugh,, 1975)
BCI	12 years +	N/A	N/A	√	X	X	√	5	10 mins	free	N/A	English, Espanol	(Fan, Flombaumb, McCandlissa, Thomas, & Michael I, 2003), (Monte, Geffen , May, & McFarland, 2010), (Melara, & Algom, 2003), (Cicerone & Azulay, 2002), (Greenwald, et al., 2001) and (Comerford, Geffen, May, Medland , & Geffen, 2002)
BrainCheck	17 years+	N/A	N/A	√	X	X	√	10	10 mins	Paid	$4.99	German, English, French, Italian, Spanish and 25 more	(Holsinger , Deveau , Boustani , & Williams , 2007), (Friedman , et al., 1994) (Jorm , Scott , & Jacomb, 1989)
ALZ	12 years +	5	N/A	√	X	X	√	100+	30mins -1 hour	Free	N/A	English	(Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000), (Tariq, Tumosa, Chibnall, Perry, & Morley, , 2006), (Tappen, Rosselli, & Engstrom, 2009), (Yesavage, et al., 1982),(Johnson, et al., 2014), (Buschke, et al., 1999)
eSLUMS	Anyone	3.9	N/A	√	X	X	X	24	7-10 mins	Free	N/A	English and Chinese	(Tariq, Tumosa, Chibnall, Perry, & Morley, , 2006)
DTRCD	12 years +	5	N/A	√	X	X	X	11	5 mins	Free	N/A	English, Czech, French, German, Portuguese, Spanish	(Torre, 2004)
Cognity	12 years +	3+	500+	√	√	√	√	30	10-15 mims	Paid	$8.99	English and Turkce	(Snyderman & Rovner, 2009) (Tariq, Tumosa, Chibnall, Perry, & Morley, , 2006)
DST	Anyone	3+	100+	√	√	√	√	30	10-15 mims	Paid	$8.49	English	No Reference available
MMSE	Anyone	3.4	5000+	X	√	X	X	10	5 mins	Free	N/A	English	(Folstein, Folstein , & McHugh, 1975)

Dementia Test	Age 60+	3+	50+	√	√	X	X	7	5 mins	Free	N/A	English	(Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002) (Ouimette & Klein , 1995)
MMSE	Anyone	3+	1000+	X	√	X	√	13	5 mins	Free	N/A	English	(Folstein, Folstein , & McHugh, 1975)
6 Cognitive Item Test (6CIT) App	12 years +	N/A	N/A	√	X	X	X	6	5 mins	Free	N/A	English	(Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002)
Cognitive Exams	Age 45+	4.4	10,000+	X	√	X	√	68	45 mins	Paid	$3.59	English	(Folstein, Folstein , & McHugh, 1975)(Yesavage, et al., 1982)

3. Discussion

3.1 DSM-5 Neurocognitive Domains Coverage

There are scientific bodies that set the guidelines for screening and diagnosis of dementia and other types of MCIs, particularly the Neurological and Communicative Disorders and Stroke & the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA), and the American Psychiatric Association (APA; McKhann et al., 1984; American Psychiatric Association, 2013). The NINCDS-ADRDA updated the diagnosis criteria of neurocognitive disorders in 2011 to enhance the medical diagnosis of AD and promote future research. The APA introduced the fifth edition of the Diagnostics and Statistics Manual (DSM-5) in 2013 by replacing the term 'dementia' with 'Major Neurocognitive Disorder' (NCD) focussing entirely on the clinical diagnosis of dementia and its types. It also addresses the severity levels of NCD and presents guidelines for clinicians to differentiate between mild and major levels of impairment along with other neurocognitive disorders falling into the same category. Hence, in the DSM-5 guidelines, major NCD refers to dementia, and mild NCD refers to other types of MCIs. DSM-5 describes the six main domains where the impairment can be present for an individual to be diagnosed as a case of major or minor NCD (American Psychiatric Association, 2013) including complex attention, executive functions, learning & memory, language, perceptual motor and social cognition.

If an individual is experiencing considerable deterioration in at least one of the above domain capabilities, a screening medical test reported by an informant, caregiver, or a clinician is required to establish the NCD symptomatology. If the problem is serious enough to interfere with the independence of the individual's daily operations, the diagnosis suggested by DSM-5 is major NCD; if the deterioration of domain capabilities is mild and has no impact on the individual's daily operations, then the diagnosis would be minor NCD (Sachdev, et al., 2014). To make a sound diagnosis, a screening medical test should be comprehensive enough to capture multiple domains described under the umbrella of DSM-5.

In this review, the comprehensiveness of a medical test refers to how well it covers the domains listed under DSM-5 for neurocognitive disorders. To understand that, each dementia screening app discussed was downloaded and evaluated with respect to the six domains of DSM-5. This becomes more complex when you look more closely at the

DSM-5 domains, and think about how performance in each domain can be measured. For example, is it really possible to separate complex attention, a term used to describe processes such as selectively paying attention only to relevant stimuli in our environments, from executive functioning, a term used to describe processes that prevent interference from distracting stimuli? Such issues were recognised by the APA task force who developed the new criteria (Ganguli et al., 2011), but the consensus was to retain these categories for diagnostic purposes. Not only are there difficulties in delineating precise definitions of each domain, but many of the tasks that can be used to measure neurocognitive performance tap into multiple domains. For example, digit span has been proposed as a measure of complex attention, with forward recall requiring repetition of items presented to the patient in the same order they are presented. This also taps into memory, and, when conducted orally, requires command of language. Moreover, because of the manipulation of items in memory demanded by the reverse digit span task (where items should be repeated in the reverse order to that in which they were presented), this variant of the task is often considered to measure working memory, which comes under the domain of executive functioning, but also requires complex attention and memory. As such, it may not always be clear from deteriorations in task performance precisely which domains are impaired. In terms of practical implications for screening, since referral only requires deterioration in one or more domains, despite issues with overlap, poor performance in any subtest will still result in referral for more in-depth assessment of patient needs. It is if some domains are not covered that there are concerns about the sensitivity of the screening tool. As a result, we took a pragmatic approach in this review, whereby some of the subtests within the apps were considered to tap into more than one domain (see Table 3 for details).

In some cases, statistical analysis has been used to demonstrate a cognitive tests ability to measure specific cognitive domains, for example Guerrero-Berroa et al (2009) examination of MMSE. However for less common tests, verification studies are not commonly available and so in such cases a test reported cognitive domain coverage has been determined based on the coverage stated by the test authors and our own discretion using the definition of each cognitive domain provided by the Johns Hopkins Psychiatry Guide (Peters & Rabins, 2017).

Table 3 illustrates the number of domains covered by each considered screening app and shows that few existing apps cover most of the cognitive domains listed under DSM-5. Many of these apps fulfil 1–4 domains, and most of the apps essentially test the user's memory. The MoCA, ALZ, BrainTest and CognitiveExams apps evaluate most of the cognitive areas defined in DSM-5 since they combine many medical tests in which each covers one or more cognitive domain(s). Therefore, these apps can be seen more comprehensive for clinical testing of MCI and AD despite the complete cognitive domains of DSM are not been evaluated. One of the notable concerns of the available apps that none of them evaluated social cognition domain. Social cognition often involves behaviours that are not socially acceptable, decisions without considering safety, and insensitivity to social standards, and not covering this domain during the screening may lead to people who are have deterioration in social cognition but were undetected. The only apps that might partly cover social cognition is ALZ and CognitiveExams using the Geriatric Depression Scale (GDS) test, which could indicates emotional behaviours by asking questions on life satisfaction, interests, perception towards life, hopefulness, happiness, problems, worries, social gatherings, etc. The DementiaScreener app also

contains the Symptoms of Dementia Screener (SDS) test, which has several questions that could indicate an issue linked with social cognition. However, the questions in the above tests would not allow for a quantifiable measure of social cognition and do not cover all areas of social cognition so a patient exhibiting a decline in social cognition may not be detected by these tests.

Table 3: Number of domains covered by each app and its tests

No	App	Sub-Test	Domains covered						No of Domains Covered
			Complex attention	Executive function	Learning and memory	Language	Perceptual motor	Social cognition	
1	MOBI-COG	Mini-Cog Dementia Screening Test	X	x	✓	x	✓	x	2
2	Brain-Test	Self-Administered Gerocognitive Exam (SAGE)	✓	✓	✓	✓	✓	x	5
3	ACE	Addenbrooke's Cognitive Examination -3 (ACE-3)	✓	x	✓	✓	✓	x	4
4	CAIDE-DRS	Memory Statement	X	x	~	x	x	x	0
5	DRT	Memory Statement	X	x	~	x	x	x	0
6	Dementia-Screener	All tests	✓	✓	✓	✓	✓	~	5
		Symptoms of Dementia Screener (SDS)	✓	✓	✓	✓	✓	~	
		AD8 Dementia Screening Interview	✓	✓	✓	x	✓	x	
7	MoCA	Montreal Cognitive Assessment (MoCA)	✓	x	✓	✓	✓	x	4
8	CDD	All tests	✓	x	✓	✓	✓	x	4
		Confusion assessment method (CAM)	~	x	~	x	~	x	
		Six item screener	X	x	✓	x	x	x	
		Abbreviated mental test (AMT)	✓	x	✓	✓	x	x	
		Montreal Cognitive Assessment (MoCA)	✓	x	✓	✓	✓	x	
		Mini-mental state examination (MMSE)	✓	x	✓	✓	✓	x	
9	BCI	All tests	✓	✓	✓	x	x	x	3
		Flanker Task	✓	✓	x	x	x	x	
		Digit Symbol Substitution Task	✓	✓	x	x	x	x	
		Stroop Task	✓	✓	x	x	x	x	
		Trail Making Test (TMT)	✓	✓	x	x	x	x	

		Recall tests	X	x	✓	x	x	x	
10	Brain-Check	All tests	X	x	✓	x	✓	x	2
		Clock Drawing	X	x	x	x	✓	x	
		7 questions from the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE)	X	x	✓	x	x	x	
11	ALZ	All tests	✓	✓	✓	✓	✓	~	5
		Mini-Cog Dementia Screening Test	X	x	✓	x	✓	x	
		Saint Louis University Mental Status Exam (SLUMS)	✓	x	✓	✓	✓	x	
		Functional Activities Questionnaire (FAQ)	✓	✓	✓	x	✓	x	
		Geriatric Depression Scale	X	x	x	x	x	~	
		Memory Impairment Screen (MIS)	X	x	✓	x	x	x	
12	eSLUMS	Saint Louis University Mental Status Exam (SLUMS)	✓	x	✓	✓	✓	x	4
13	DTRCD	Dementia Risk Calculator	✓	x	✓	✓	✓	x	4
14	Cognity	All tests	X	x	✓	x	✓	x	2
		Clock Drawing	X	x	x	x	✓	x	
		Mental Status Examination	X	x	✓	x	x	x	
15	DST	Dementia Screening Test	✓	✓	✓	✓	x	x	4
16	MMSE	Mini-mental state examination (MMSE)	✓	x	✓	✓	✓	x	4
17	Dementia Test	6-item Cognitive Impairment Test (6CIT)	✓	x	✓	x	x	x	2
18	MMSE	Mini-mental state examination (MMSE)	✓	x	✓	✓	✓	x	4
19	6CIT	6-item Cognitive Impairment Test (6CIT)	✓	x	✓	x	x	x	2
20	Cognitive Exams	All tests	✓	x	✓	✓	✓	~	4
		Mini-mental state examination (MMSE)	✓	x	✓	✓	✓	x	
		Clock Drawing	X	x	x	x	✓	x	
		Geriatric Depression Scale	X	x	x	x	x	~	
		Verbal Fluency Test	X	x	x	✓	x	x	
		Katz Test	X	x	x	x	~	x	

~ denotes Marginal/Partial domain coverage.

3.2 Validity and Reporting

Validity is a significant criterion that defines the suitability of a screening app for clinical and primary care settings. This section evaluates the validity of all the considered apps to determine the ones that are suitable to be used in a clinical setting. Table 4 lists all the medical tests used in considered apps along with their validity measures. The table shows two types of validated tools: those that are recommended by an accepted international body, and others that are validated through an exclusive experiment carried out in a clinical setting. However, most of the medical tests considered are validated through one or both methods. BrainCheck's patients' questions, and the DST app are not validated through medical body recommendation. Nevertheless, the complete BrainCheck app was validated in a study by Yang et al., (2017) using 30 patients, and 568 controls and showed acceptable performance (83% sensitivity and 87% specificity).

MCI and dementia medical tests including MMSE, Mini-COG, AD8 Dementia Screening Interview, 7-item version of the IQCODE, MIS, eSLUM, MoCA, e-SAGE, and GPCOG are recommended by the Alzheimer's Association and American Academy of Neurology (AAN). This means MoCA, MMSE, ALZ, MOBI-COG, eSLUM and BrainCheck can be considered validated apps that can be carried out in clinical settings by clinicians. According to the Alzheimer's Association's guidelines, a patient should be assessed for the health risk and possible symptomatology before the annual wellness visit which can possibly occur in the primary care setting. Thus, most of the apps that are brief (less than 10 minutes), can be administered in clinical setting according to the Alzheimer's Association and can be considered as valid screening tools for cognitive assessment (Cordella et al., 2013). Further, the performance of these tools is validated in terms of sensitivity and specificity through several research studies (Foster, et al., 2019; Arevalo-Rodriguez et al., 2015; Borson, Scanlan, Watanabe, Tu, & Lessig, 2006; Kansagara & Freeman, 2010; Brodaty et al., 2002).

Reporting corresponds to providing meaningful results electronically to aid the clinician's formal diagnosis. Out of all the considered apps, MOBI-COG, DRT, DementiaScreener, DTRCD, DementiaTest, and 6CIT generate reports with scores. Sometimes, the score presented is briefly interpreted with basic recommendations such as the need for further medical assistance. Apps such as BrainTest®, ACE, MoCA, BCI, eSLUMS, Cognity, DST, ALZ and CognitiveExams produce a formal medical report in pdf format that can be printed and presented to the clinician during a medical follow up. Some apps have the option to share the reports immediately via e-mail or other electronic means. Most of the apps track the results history, and a few apps such as BCI and BrainTest® are capable of recording and tracking the results and progress of multiple patients. BrainTest® does not provide interpretation of the score immediately or generate a report, however, results interpreted by a physician within an instructional video will be e-mailed to the user within five working days using the BrainTest® app.

Usability, also known as user friendliness can be defined as a tool's ability to achieve specified goal in an effective, efficient and satisfactory manner with in a specified context (Bolchini, Finkelstein, Perrone, & Nagl, 2009; Douglas

and et al, 2011). But the question is what makes a tool effective and efficient in achieving the specified goals. When it comes to biomedical science dementia screening tools, usability is an important parameter can be used to evaluate their performance. Therefore, each tool needs to be followed by an exclusive usability study to demonstrate their user-friendliness, effectiveness, efficiency, and accessibility. But only few of the above discussed tools have usability studies carried out to test their user friendliness for both patient (case) and the clinician.

Newman et al., (2018) used trainee clinical psychologists and post graduate health science students as the sample population to administer the ACE app to evaluate the usability of the tool. The findings suggested that the ACE app had a considerable error in items pertaining to the language domain. An adjustment was made to remove the incorrect, confusing naming to enhance the usability of the app in a clinical setting. Similarly, Nirjon et al., (2014) conducted a study on eight healthy individuals (four males and four females) to demonstrate that the tool can perform all three tasks of remembering words, clock drawing, and recalling words effectively in a standard primary care setting. The authors used a separate questionnaire with a score range from 1–10 (10 indicates high level of usability and 0 indicates low level of usability) given to participants to evaluate the usability of the MOBI-COG app. However, the findings of the study indicated that the tool is suitable for dementia screening in both clinical and primary care settings.

Table 4: Validity details of the considered apps

No	App	Medical Examination	Validity		
			Clinically Validated	Recommended By	Reference
1	MOBI-COG	Mini-COG dementia screening test	✓	The Alzheimer's Association	(Cordella, et al., 2013)
2	BrainTest	Self-Administrated Gerocognitve Exam (e-SAGE)	✓	American Academy of Neurology (AAN)	(Foster, et al., 2019)
3	ACE	Addenbrooke's Cognitive Examination (ACE-III)	✓	Department of Health and the Alzheimer's Society in the UK.	(Bruno & Schurmann , 2018)
4	CAIDE-DRS	Cardiovascular Risk factors, Aging and Incidents of Dementia risk score(CAIDE) risk score	✓	_	(Sindi, et al., 2015)
5	DementiaScreener	Symptoms of Dementia Screener (SDS)	✓	_	(Flaherty, Midden, & Mast, 2019)
		AD8 Dementia Screening Interview	✓	The Alzheimer's Association	(Cordella, et al., 2013)
6	MoCA	Montreal cognitive assessment	✓	American Academy of Neurology (AAN)	(Foster, et al., 2019)
7	CDD	Confusion assesment method (CAM)	✓	_	(Waszynsk, 2012)
		Six item screener	✓	_	(Carpenter, DesPain, Keeling, Shah, & Rothenberger, 2011)
		Abbriviated mental test (AMT)	✓	_	(Jitapunkul, Pillay, & Ebrahim, 1991)
8	BCI	Flanker Task The Digit Symbol Substitution Task	✓	Registered as a Class II medical device with	(BrainCheck Inc, 2019)

		The Stroop Task		the U.S. Food and Drug Administration (FDA)	
		The Trail Making Test (TMT)			
		Balance and coordination			
		The Immediate and Delayed Recall Tests			
9	BrainCheck	3 patient questions (BrainCheck)	X	X	X
		Clock drawing test (CDT)	✓	the National Collaborating Centre for Mental Health (UK)	(National Collaborating Centre for Mental Health , 2007)
		7-item version of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE)	✓	The Alzheimer's Association	(Cordella, et al., 2013)
10	ALZ	Functional Activities Questionnaire (FAQ)	✓	National Alzheimer's Coordinating Center (NACC)	(Mayo, 2016)
		Hachinski Ischemic Score (HIS)	X	–	(Gay, Taylor, Hohl, Tolnay, & Staehelin, 2008)
		Memory Impairment Screen (MIS)	✓	American Academy of Neurology (AAN)	(Foster, et al., 2019)
		General practitioner assessment of cognition (GPCOG)	✓	The Alzheimer's Association	(Cordella, et al., 2013)
11	eSLUMS	Saint Louis University Mental Status (SLUMS) Exam	✓	American Academy of Neurology (AAN)	(Foster, et al., 2019)
12	DTRCD	Mental Status Examination (MSE) tool	✓	U.S. Preventive Services Task Force	(Snyderman & Rovner, 2009)
13	DST	Dementia Screening Test	X	X	X
14	MMSE	Mini Mental State Examination (MMSE)	✓	American Academy of Neurology (AAN)	(Foster, et al., 2019)
15	DementiaTest	6CIT	✓	Alzheimer's Society and the National Collaborating Centre for Mental Health (UK)	(Ballard, Burns, Corbett , Livingston, & Rasmussen, 2013)
		Structured Clinical Interview (SCIDS)	✓	American Society of Addiction Medicine	(Gerdner , Kestenberg , & Edvinsson , 2014)
16	CognitiveExams	Geriatric Depression Scale	✓	–	(Mitchell , Bird , Rizzo, & Meader , 2011)
		Katz basic activities of daily living as functional assessment	✓	The Hartford Institute for Geriatric Nursing, New York University Rory Meyers College of Nursing	(McCab, 2019)
		Walk Test of 6 minutes	✓	Neurology section of American Physical Therapy Association's Multiple Sclerosis Taskforce (MSEDGE), Parkinson's Taskforce (PD EDGE), Spinal cord injury taskforce, stroke taskforce, Traumatic brain injury taskforce and Vestibular taskforce	(AbilityLab, 2019)
		Fluency Test	✓	_	(Herrera-García, et al., 2019)

3.3 Performance

There are many measures that can be used to evaluate the performance of a medical screening tool, including sensitivity and specificity, which are widely used for this purpose. These measures define how capable a screening tool is of recognizing dementia-related symptomology and the degree to which it can distinguish dementia from other MCIs (Maxim, Niebo, & Utel, 2014). Sensitivity defines the tool's ability to classify an individual with the disease as positive; specificity defines the tool's ability to classify a person without disease as negative (Goetzinger & Odibo, 2011).

Several research studies have been conducted to evaluate the performance of dementia-screening methods like MMSE, Mini-COG, eSLUMS, CPCOG, and MIS among others for different populations (Borson et al., 2006; Cordella et al., 2013; Arevalo-Rodrigue et al., 2015; Foster, et al., 2019; Kansagara & Freeman, 2010; Tsoi, Chan, Hirai, Wong, & Kwok, 2015). In this research, reported sensitivity and specificity figures are considered using multiple research studies. Table 5 summarizes performance measures along with the associated study from which these values are obtained. According to the information in Table 5, the Mini-COG assessment, ACE-III, CAM, and eSLUMS assessments are the highest performing tools with both sensitivity and specificity values over 90% (at least on the sample instances used in their corresponding research studies). Generally, it is not the norm to achieve very high values in both sensitivity and specificity simultaneously. The trade-off value between both measures is often considered as the best.

Table 5: Reported sensitivities and specificities of medical exams used by the app

App	Medical examination	Medical exam reference	Sensitivity	Specificity	Study performance reference
MOBI-COG	Mini-COG dementia screening test	(Borson, Scanlan, Brush, Vitaliano, & Dokmak, 2000)	95.00-100.00%	92.00-98.00%	(Borson S. , Scanlan , Watanabe , Tu , & Lessig , 2006)
BrainTest	Self-Administrated Gerocognitve Exam (e-SAGE)	(Scharre D. , et al., 2010)	71.00%	90.00%	(Scharre D. , Chang, Nagaraja, Vrettos, & Bornstein, 2017)
ACE	Addenbrooke's Cognitive Examination (ACE-III)	(Noone, 2015)	93.00-100.00%	96.00-100.00%	(Bruno & Vignaga, 2019)
CAIDE-DRS	Cardiovascular Risk factors, Aging and Incidents of Dementia risk score(CAIDE) risk score	(Kivipelto, et al., 2006)	77.00%	63.00%	(Butcher, 2007)
DementiaScreener	Symptoms of Dementia Screener (SDS)	(Flaherty, Midden, & Mast, 2019)	78.40%	84.00%	(Flaherty, Midden, & Mast, 2019)
	AD8 Dementia Screening Interview	(Galvin , et al., 2005)			
MoCA	Montreal cognitive assessment	(Ziad S. Nasreddine, et al., 2005)	81.00%	86.00%	(Julayanont, et al., 2015)
CDD	Confusion assesment method (CAM)	(Inouye, et al., 1990)	94.00-100.00%	89.00-95.00%	(Waszynsk, 2012)

	Six item screener	(Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002)	68.00–80.00%	74.00–80.00%	(Carpenter, DesPain, Keeling, Shah, & Rothenberger, 2011)
	Abbriviated mental test (AMT)	(Hodkinson H. , 1972)	70.00–80.00%	74.00–90.00%	(Jitapunkul, Pillay, & Ebrahim, 1991)
BCI	Flanker Task	(Fan, Flombaumb, McCandlissa, Thomas, & Michael I, 2003)	81.00%	94.00%	(Groppell, et al., 2019)
	The Digit Symbol Substitution Task	(Monte, Geffen , May, & McFarland, 2010)			
	The Stroop Task	(Melara, & Algom, 2003)			
	The Trail Making Test (TMT)	(Cicerone & Azulay, 2002)			
	Balance and coordination	(Greenwald, et al., 2001)			
	The Immediate and Delayed Recall Tests	(Comerford, Geffen, May, Medland , & Geffen, 2002)			
BrainCheck	3 patient questions (BrainCheck)	(Holsinger , Deveau , Boustani , & Williams , 2007)	85.80%	74.30%	Ehrensperger, et al., (2014)
	Clock drawing test (CDT)	(Friedman , et al., 1994)			
	7-item version of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE)	(Jorm , Scott , & Jacomb, 1989)	81.40%	75.70%	
ALZ	Functional Activities Questionnaire (FAQ)	(Tappen, Rosselli, & Engstrom, 2009),	80.30%	87.00%	(Teng, et al., 2011)
	Hachinski Ischemic Score (HIS)	(Johnson, et al., 2014),	89.00%	89.30%	(Moroney , et al., 1997)
	Memory Impairment Screen (MIS)	(Buschke, et al., 1999)	68.00-86.00%	84.00-96.00%	(Tsoi, Chan, Hirai, Wong, & Kwok, 2015)
	General practitioner assessment of cognition (GPCOG)	(Brodaty , et al., 2002)	85.00%	86.00%	(Brodaty , et al., 2002)
eSLUMS	Saint Louis University Mental Status (SLUMS) Exam	Saint Louis University Mental Status (SLUMS)	98.00-100.00%	98.00-100.00%	(Kansagara & Freeman , 2010)
Cognity	Mental Status Examination (MSE) tool	(Snyderman & Rovner, 2009)	71.00-92.00%	52.00-96.00%	(Snyderman & Rovner, 2009)
DST	Dementia Screening Test	Dr. S. Horn (2019)	96.00%	N/A	Developer's note on Appstore
MMSE	Mini Mental State Examination (MMSE)	(Folstein, Folstein, & McHugh, 1975)	23.00-76.00%	40.00-94.00%	(Arevalo-Rodriguez, et al., 2015)
DementiaTest	6-item Cognitive Impairment Test (6CIT)	(Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002)	78.00-90.00%	100%	(Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002)
	Structured Clinical Interview (SCIDS)	(Ouimette & Klein, 1995)	75.00-100.00%	100%	(Gerdner , Kestenberg , & Edvinsson , 2014)
CognitiveExams	Geriatric Depression Scale	(Yesavage, et al., 1982)	92.50%	77.20%	(Mitchell , Bird , Rizzo, & Meader , 2011)
	Katz basic activities of daily living as functional assessment	(Aske, 1990)	38.00%	N/A	(Hartigan , 2006)
	Walk Test of 6 minutes	(Abbott, et al., 2004)	82.00%	84.00%	(AbilityLab, 2019)
	Fluency Test	(Caramelli, Carthery-Goulart, Porto,	N/A	N/A	N/A

		Charchat-Fichman, & Nitrini, 2007).			

3.4 The Role of Intelligent Methods

ML is a hot topic that has a direct impact on prognosis of dementia and other related MCIs. ML is a process that allows medical systems using search techniques to automatically learn from data to improve the accuracy and efficiency of the screening with little human involvement (Gupta & Katarya, 2020; Sammut & Webb, 2017). The potential of these technologies in dementia prognosis is vast, yet few studies have been carried out to discover way to integrate them with conventional assessment tests (Chua, et al., 2019; Shankle, Mani, Pazzani, & Smyth, 2005; Bennasar, Setchi, & Hicks, 2014). The accuracy of AD screening tools is paramount as it ensures that no patient is left undiagnosed, and it also facilitates the speedy implementation of treatment plans. Hence, assessment tools should be automated to yield better performance and accuracy and to offer rich knowledge for clinicians to make diagnostic-related decisions. Further, patients involved are often elderly and likely to have other physical, social, and accessibility issues, therefore utilizing technology for screening instruments makes the entire process simple, accurate, and convenient for them and the clinicians and caregivers.

Out of all considered mobile screening apps, few employ AI in dementia screening. According to Snyderman and Rovner, (2009), Cognity is the first mobile app to utilize AI for dementia screening; this app analyses a photo of a clock drawing done by the user and then, based a large sample of patients, determines the probability of dementia risk. The ACE app also has an inbuilt mechanism for automated administration, scoring, and reporting using the human factor approach. Findings of Newman et al., (2018) suggest that the computerized version (ACE app) can capture measures more accurately than the original version of the ACE medical assessment. MOBI-COG, the digitalized version of Mini-Cog test, is also a fully automated mobile app that uses ML techniques to recognize hand-written digits and characters. It uses a k-NN classifier (Gonzalez & Woods, 2002) to identify the characteristics of the clock drawing done by the suspected individual. Findings of Nirjon et al., (2014) indicate that the automated mobile version is more capable of evaluating the correctness of a clock drawing to detect the presence of dementia and other types of MCIs than the conventional pen-and-paper based Mini-Cog test, with an accuracy of 99.5%.

4. Future Works and Limitations

Although many studies have been carried out on the use of ML to predict the risk of dementia, few have focused on the use of ML techniques to enhance the performance of traditional screening examinations. For instance, Okada, Inoue, Imai, and Noguchi, (2019) proposed a novel approach that uses a humanoid robot to capture the interactive behaviours, and ubiquitous sensors to capture the daily activities of the suspected individual to predict the score of dementia scale. Chiu et al., (2019) presented a new questionnaire named NMD-12 (New ML-derived screening instrument) with 45 items developed using the Information Gain (IG) feature selection method (Quinlan, 1993) to

differentiate between normal cognition, MCI, and dementia. Studies such as Chua et al., (2019) highlight the role of virtual reality (VR) technology in cognitive impairment screening. The study measures the ability of VR-based screening technology in capturing learning, memory, perceptual-motor function, and the executive function of the suspected individual against the performance of a few traditional tools like MoCA, AMT, and MMSE. The results of the study pinpointed that there is a high feasibility for VR-based screening tools over conventional methods in evaluating cognitive functions of the elderly.

Deep learning which is an advancement of AI can provide diagnosis and screening methods of dementia with advanced capabilities (An, et al., 2020; Moreno-Garc´ıa, Elyan, and Jayne, 2018). For example, with the use of deep learning technique not only clinicians are able utilize a powerful classification systems with high predictive accuracy but also these techniques can offer the impactful dementia features and other useful information to the clinicians to aid the diagnostic process. Choi et al., (2018) designed a two-phase cost-effective classification system based on deep neural networks (DNNs) that model neuropsychological assessment scores of MMSE and KLOSCAD Neuropsychological Assessment Battery for dementia screening. There is no need to have a separate phase for feature extraction or selection within deep learning techniques as in MI. More importantly, deep learning methods process unstructured and sophisticated features such as video, audio, and images in real time, providing diagnosticians and with enhanced dementia computed aided methods that guarantee improved performance empowering medical professionals.

Integrating technological advancements into traditional dementia screening assessments is a promising topic that has not yet gained enough attention from the relevant research bodies. Therefore, more research that focuses on improving the accuracy and efficiency of the existing dementia screening tools using advanced intelligent technology is required. As limited research has been carried out on the adaptation of ML techniques that can be integrated into these mobile apps to enhance their functionalities, additional research is recommended on the use of intelligent algorithms in dementia screening and diagnosis. Replacing conventional scoring functions used in traditional medical screening and diagnosis methods with more advanced intelligent classification systems will be advantageous for clinicians. This can be attributed to the fact that these methods are less biased to the final outcomes as they are learnt from historical datasets (former controls and cases) and can be exploited by the clinician to aid the decision-making process.

There are a large number of cognitive diagnostic tests related to dementia and MCI that aim to measure a patient's ability in distinct cognitive areas. While these cognitive tests have been utilized by clinicians to demonstrate the presence of dementia or MCI, there is no clear consensus as to where these procedures fit into the process of dementia diagnosis (Sachdev et al., 2014; Baldwin, & Farias, 2009). Therefore, other research that can critically analyse an array of commonly employed cognitive diagnostic tests to determine how they can be used specifically within the DSM-5 criteria for AD diagnosis, is advantageous.

One limitation surrounds the conceptual overlap of the DSM neurocognitive domains that we already discuss in section 3.1, and how because of task impurity patient scores on individual tasks may not correspond directly to DSM domains. Another limitation that applies to this review and others is that while we have accumulated the evidence available, in a systematic manner, your conclusions are only as good as the research you have included within your review; e.g. sensitivity/specificity data are based on studies done in other laboratories.

This research focuses on evaluating existing dementia screening apps based on a few parameters. It doesn't address the technical aspects of these tools (user friendliness, attractiveness, convenience of user interfaces, use of audio and video means, understandability of items, and ease of administration); further research could be carried out to evaluate the technicalities of these tools as well as the domains of human factors (usability, accessibility, literacy) aspects of the dementia mobile apps.

5. Conclusions

Although there are many apps available to screen dementia, many of them are just partly covering the neurocognitive domains defined in the DSM-5 gold standard. There is also a lack of systematic research to evaluate these apps in terms of performance, validity, and adaptation of intelligent techniques such as ML and AI among others. Therefore, to fill the gap, this research reviews and critically analyses existing dementia and MCI screening apps available on Android and Apple Stores based on inclusion and exclusion criteria. Five criteria: DSM-5 Coverage, Performance, Validity, and Reporting, and the Use of ML and AI techniques are discussed for 20 different screening apps which have been filleted from 275 available apps. More importantly, this research proposed a mapping approach for the available screening apps based on the defined six cognitive domains for neurological disorders within DSM-5. In particular, we identified the possible medical tests available within the dementia screening apps and then created a mapping model which links each medical test (based on its activities/ questions) and the appropriate cognitive area(s) it covered, which indeed can provide advantageous information to the diagnosticians of which dementia screening apps cover wider numbers of cognitive domains.

The findings suggest MoCA, ALZ and CognitiveExam apps cover most of the cognitive domains defined in DSM-5 for neurocognitive disorders, making them suitable for comprehensive medical diagnosis within clinical setup by diagnosticians. Most of the screening apps used such as IQCODE, MIS, SLUMS, e-SAGE, and GPCOG, are validated through independent research studies and have been utilized by medical bodies. One concerning cognitive area that the considered apps have not covered or at least partially covered is social cognition, therefore there is a need to include tests or questions related to social cognition while screening for MCI and AD using mobile tools. There is evidence that ML and AI techniques have been utilized in dementia screening tools to enhance outcome accuracy and to accelerate the entire screening process, for example, in apps such as Cognity. Cognity uses AI to evaluate the characteristics of a clock drawing by the user to screen for dementia. The ACE-Mobile and MOBI-COG apps employ auto scoring, reporting mechanisms, and ML classifiers such as k-NN to reduce the error of the screening process.

The findings of the review suggested that most of the considered apps can be used in both primary and clinical settings. Although usability studies are a standard in neurocognitive screening, the ACE-Mobile and MOBI-COG apps alone have usability studies to confirm their applicability in real-world settings.

The conducted research helps patients, caregivers, families, healthcare professionals including clinicians, identify the advantages and shortcomings of the screening tools for dementia and MCI available on Google Play and Apple App stores.

References

Abbott, R., White,, L., Ross, G., Masaki, K., Curb,, J., & Petrovitch, H. (2004). Walking and Dementia in Physically Capable Elderly Men. *American Medical Association*, 1447-1454.

An N., Ding H. and Yang J., Au R., Ang T.F.A. (2020) Deep ensemble learning for Alzheimers disease classification (2020) Journal of Biomedical Informatics, Volume 105, May 2020, 103411.

Ballard, C., Burns, A., Corbett , A., Livingston, G., & Rasmussen, J. (2013). *Helping you to assess cognition - A practical toolkit for clinicians.* London E1W 1LB: Alzheimer's Society Research.

Brodaty, H., & Gibson, L. (2006). What Is the Best Dementia Screening Instrument for General Practitioners to Use? *Brodaty, H., Low, L.-F., Gibson, L., & Burns, K. (2006). What Is the Best Dementia Screening Instrument for GenThe American Journal of Geriatric Psychiatry, 14(5),* , 391–400. doi:10.1097/01.jgp.0000216181.20416.b2 .

Douglas, C. &.-G. (2011). Socio-Cultural Characteristics of Usability of Bioinformatics Databases and Tools. Interdisciplinary Science Reviews. 36, 55-71.

Ernst , R., & Hay , J. (1999). The u.s. economic and social costs of alzheimer's disease revisited. *American Journal of Public Health*, 1261-1264.

Hegde, V. (2013). Role of human factors/usability engineering in medical device design. . *Reliab Maintainab Symphosium (RAMS)*.

Inouye, S., Dyck, C., Alessi, C., Balkin, S., Siegal, A., & Horwitz, R. (1990). Clarifying Confusion: The Confusion Assessment Method - A New Method for Detection of Delirium . *Journal of Annals of Internal Medicine*, 941-948.

Klimova, B. (2017). Mobile Phone Apps in the Management and Assessment of Mild Cognitive Impairment and/or Mild-to-Moderate Dementia: An Opinion Article on Recent Findings. *Frontiers in Human Neuroscience*, doi: 10.3389/fnhum.2017.00461.

McCleary , R., Shankle , W., Mulnard , R., & Dick , M. (1996). Ishihara test performance and dementia. *Journal of the Neurological Sciences*.

Nabney, I. (2004). Efficient training of RBF networks for classification. . *International Journal of Neural Systems* , 201-208.

Noone, P. (2015). Addenbrooke's Cognitive Examination-III. *Occupational Medicine, Volume 65, Issue 5*, 418–420.

Pang , G.-H., & Kwong , E. (2015). Considerations and Design on Apps for Elderly with . *Mild-to-moderate Dementia*, doi:10.1109/icoin.2015.7057910.

Peter , C. (2007). : A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality. . *Statistics in Medicine* , 2937-2957. .

AbilityLab. (2019). *6 minutes Walk Test*. Retrieved from AbilityLab; Rehabilitation measure database: https://www.sralab.org/rehabilitation-measures/6-minute-walk-test

Alzheimer's Association. (2019). Alzheimer's Disease Facts and Figures. *Alzheimers Dementia*, 321-87.

Alzheimer's Association. (2019). *New diagnostic criteria and guidelines for Alzheimer's disease*. Retrieved from Alzheimer's Association: https://www.alz.org/research

American Psychiatric Association. (2019). *Diagnostic and Statistical Manual of Mental Disorders (DSM–5)*. Retrieved 11 17, 2019, from American Psychiatric Association: https://www.psychiatry.org/psychiatrists/practice/dsm

Apple Inc. (2019). *6CIT*. Retrieved from App Store Previ: https://apps.apple.com/us/app/6cit/id1442965681

Apple Inc. (2019, 10 26). *Dementia Test - risk calculator of dementia*. Retrieved from App Store Preview: https://apps.apple.com/us/app/dementia-test-risk-calculator-of-dementia/id1014958634

Apple Inc. (2019). *MoCA App*. Retrieved from App store Preview: https://apps.apple.com/us/app/moca-app/id1206246590

Arevalo-Rodriguez, I., Smailagic, N., Figuls, , M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., . . . Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database System Review*, doi: 10.1002/14651858.CD010783.pub2.

Arevalo-Rodriguez, I., Smailagic, N., Roqué i Figuls , M., Ciapponi , A., Sanchez-Perez , E., Giannakou , A., . . . Cullum , S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI) (Review). *The Cochrane Library*, 1-62.

Armstrong, N., Nugent, C., Moore , G., & Finlay, D. (2010). Using smartphones to address the needs of persons with Alzheimer's disease. *Armstrong, N., Nugent, C., Moore, G., & Finlay, D. (2010). Using smartphones to address the needs of persons with Alzheimer's disease. Annals of Telecommunications - Annales Des Télécommunications, 65(9-10),* , 485–495. doi:10.1007/s12243-010-0165-3 .

Aske, D. (1990). The Correlation Between Mini-Mental State Examination Scores and Katz ADL Status Among Dementia Patients. *Rehabilitation Nursing*, 140-146.

Baldwin, S., & Farias, S. T. (2009). Unit 10.3: Assessment of Cognitive Impairments in the Diagnosis of Alzheimer's Disease". Current protocols in neuroscience/editorial board, Jacqueline N. Crawley...[et al.], Unit10-3.

Bennasar, M., Setchi, R., & Hicks , Y. (2014). Cascade Classification for Diagnosing Dementia. *IEEE International Conference on Systems, Man, and Cybernetics* , (pp. 2535-2540). San Diego, CA, USA.

Berauk, V., Murugiah, M., Soh, Y., Sheng, Y., Wong, T., & Ming, L. (2017). Mobile Health Applications for Caring of Older People: Review and Comparison. *Therapeutic Innovation & Regulatory Science* , 1-9.

Bolchini, D., Finkelstein, A., Perrone, V., & Nagl, S. (2009). Better bioinformatics through usability analysis. Bioinformatics, 406–412.

Borson , S., Scanlan , J., Watanabe , J., Tu , S., & Lessig , M. (2006). Improving identification of cognitive impairment in primary care. *International Journal of Geriatric psychiatry*, 349-355.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The Mini-Cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry, 15(11),* , 1021–1027.

BrainCheck. (2019). *About Us*. Retrieved from BrainCheck: https://braincheck.com/about

BrainTest Inc. (2013). *Take the brain test* . Retrieved from Brain test: https://braintest.com/

Brieman , L., Friedman , J., Olshen , R., & Stone , C. (1984). *Classification and Regression Trees.* Belmont: Wadsworth.

Brodaty , H., Pond , D., Kemp , N., Luscombe , G., Harding , L., Berman , K., & Huppert , F. (2002). The GPCOG: a new screening test for dementia designed for general practice. *Journal of the American Geriatric Society*, DOI: 10.1046/j.1532-5415.2002.50122.x.

Brown, H., D'Amico, F., Knapp, M., Orrell, M., Rehill, A., Robinson, L., & Vale, L. (2019). A cost effectiveness analysis of maintenance cognitive stimulation therapy (MCST) for people with dementia: examining the influence of cognitive ability and living arrangements. *Aging and Mental health*, 602-607.

Bruno , D., & Schurmann , V. (2018). Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. *Journal of Neuropsychiatric Disease and Treatment* , 441—447.

Bruno, D., & Vignaga, S. (2019). Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. *Neuropsychiatric Disease and Treatment*, 441–447.

Buschke, H., Kuslansky, G., Katz, M., Stewart, W., Sliwinski, M., Eckholdt, H., & Lipton, R. (1999). Screening for dementia with the Memory Impairment Screen. *Neurology*, DOI: https://doi.org/10.1212/WNL.52.2.231.

Butcher, J. (2007). CAIDE Dementia Risk Score Validated in Study. *Clinical Neurology News*, 22a-22b.

Callahan, C., Unverzagt, F., Hui, S., Perkins, A., & Hendrie, H. (2002). *Six-Item Screener to Identify Cognitive Impairment among Potential Subjects for Clinical Research.* Indiana: Lippincott Williams & Wilkins.

Caramelli, P., Carthery-Goulart, M., Porto, C., Charchat-Fichman, H., & Nitrini, R. (2007). Category Fluency as a Screening Test for Alzheimer Disease in Illiterate and Literate Patients. *Alzheimer Disease & Associated Disorders*, 65–67.

Carpenter, C., DesPain, B., Keeling, T., Shah, M., & Rothenberger, M. (2011). The Six-Item Screener and AD8 for the Detection of Cognitive Impairment in Geriatric Emergency Department Patients. *Annals of Emmergency Medicine*, 653–661.

Chertkow, H. (2008). Introducing a series based on the Third Canadian Consensus Conference on the. *Canadian Medical Association or its licensors*, 316-321.

Chewy Logic, LLC. (2019). *A digital version of St Louis University Mental Status Exam (SLUMS)*. Retrieved from eSLUMS: https://www.eslumstest.com/

Choi, H., Choe, J., Kim, H. et al. Deep learning based low-cost high-accuracy diagnostic framework for dementia using comprehensive neuropsychological assessment profiles. BMC Geriatr 18, 234 (2018). https://doi.org/10.1186/s12877-018-0915-z

Chiu, P.-Y., Tang, H., Wei, C.-Y., Zhang, C., Hung, G.-U., & Zhou, W. (2019). NMD-12: A new machine-learning derived screening instrument to detect mild cognitive impairment and dementia. *PLoS ONE 14(3)*, e0213430. https://doi.org/10.1371/journal.pone.0213430.

Chua, S., Tan, N., Wong, W., Allen Jr, J., Quah, J., Malhotra, R., & Østbye, T. (2019). Virtual Reality for Screening of Cognitive Function in Older Persons: Comparative Study. *Journal of medical internet Reserach* , DOI: 10.2196/14821.

Cicerone , K., & Azulay, J. (2002). Diagnostic Utility of Attention Measures in Postconcussion Syndrome. *Journal The Clinical Neuropsychologist*, 280-289. Retrieved from https://doi.org/10.1076/clin.16.3.280.13849

Comerford, V. E., Geffen, G. M., May, C., Medland , S., & Geffen, L. (2002). A Rapid Screen of the Severity of Mild Traumatic Brain Injury. *Journal of Clinical and Experimental Neuropsychology*, 409-419. Retrieved from https://doi.org/10.1076/jcen.24.4.409.1044

Cordella, C., Borson, S., Boustan, M., Chodosh, J., Reuben, D., Verghese, J., . . . Fried, L. (2013). Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimer's & Dementia 9* , 141–150.

Dallora, A., Eivazzadeh, S., Mendes, E., Berglund, J., & Anderberga, P. (2016). Prognosis of dementia employing machine learning and microsimulation techniques: a systematic literature review. *Procedia Computer Science 100* , 480 – 488 .

Ehrensperger, M., Taylor, K., Berres, M., Foldi, N., Dellenbach, M., Bopp, I., . . . Monsch, A. (2014). BrainCheck – a very brief tool to detect incipient cognitive decline: optimized case-finding combining patient- and informant-based data. *Alzheimer's Research & Therapy*, doi: 10.1186/s13195-014-0069-y.

Fan, J., Flombaumb, J., McCandlissa, B., Thomas, K., & Michael I, .. (2003). Cognitive and Brain Consequences of Conflict. *NeuroImage*, 42-57.

Flaherty, L., Midden, A., & Mast, B. (2019). Psychometric Evaluation of the Symptoms of Dementia Screener (SDS) in a Geriatric Primary Care Population. *Journal Clinical Gerontologist*, 504-511. doi:DOI: 10.1080/07317115.2018.1453906

Folstcin, M., Folstein, S., & McHugh,, P. (1975). "Mini-Mental State" a practical method for grading the cognitive state of patients for the clinician. *Psychiatric Research, 12,* , 189-198.

Folstein, M., Folstein , S., & McHugh, P. (1975). Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric, 12 (3)*, 189-198,.

Foster, N., Bondi, M., Das, R., Foss, M., Hershey, L., Koh, S., . . . Wang, D. (2019). Quality improvement in neurology Mild cognitive impairment quality measurement set. *Neurology*, DOI: https://doi.org/10.1212/WNL.0000000000008259.

Friedman , M., Leach , L., Kaplan , E., Winocur , G., Shulman , K., & Delis, D. (1994). *DC: Clock Drawing: A Neuropsychological Analysis.* New York: Oxford University Press.

Ganguli, M., Blacker, D., Blazer, D. G., Grant, I., Jeste, D. V., Paulsen, J. S., Petersen, R. C., & Sachdev, P. S. (2011). Classification of neurocognitive disorders in DSM-5: a work in progress. The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry, 19(3), 205–210. https://doi.org/10.1097/jgp.0b013e3182051ab4

Gay, B., Taylor, K., Hohl, U., Tolnay, M., & Staehelin, H. (2008). The validity of clinical diagnoses of dementia in a group of consecutively autopsied memory clinic patients. *The Journal of Nutrition Health and Aging, 12(2), 132–137. doi:10.1007*, 132–137.

Gerdner , A., Kestenberg , J., & Edvinsson , M. (2014). Validity of the Swedish SCID and ADDIS diagnostic interviews for substance use disorders: Sensitivity and specifi city compared with a LEAD golden standard. *Nordic Journal of Psychiatry*, 1 – 9.

Goetzinger , K., & Odibo, A. (2011). Statistical analysis and interpretation of prenatal diagnostic imaging studies, Part 1: evaluating the efficiency of screening and diagnostic tests. *Journal of Ultrasound Medicine*, 1121-1127.

Gonzalez, R., & Woods, R. (2002). *Digital image processing, 2nd edition.* Upper Saddle River, New Jersey: Prentice Hall.

Google. (2016). *MMSE*. Retrieved from Google Play: https://play.google.com/store/apps/details?id=com.yasintanriverdi.mmse

Google. (2017). *Dementia & Alzheimer's Memory Diagnosis Test: MMSE* . Retrieved from Google Play: https://play.google.com/store/apps/details?id=com.alzheimers_mme

Google. (2019). *Cognitive Exams*. Retrieved from Google Play: https://play.google.com/store/apps/details?id=br.com.digos.examescognitivos

Google. (2019). *Dementia Risk Tool*. Retrieved from Google Play: https://play.google.com/store/apps/details?id=com.dementiarisktool

Google. (2019). *DST - Dementia Screening Test, Alzheimer Test*. Retrieved from Google Play: https://play.google.com/store/apps/details?id=com.dementiascreeningtest

Gupta A. Katarya R. (2020) Social media based surveillance systems for healthcare using machine learning: A systematic review. Journal of Biomedical Informatics. Volume 108, August 2020, 103500.

Guerrero-Berroa, E., Luo, X., Schmeidler, J., Rapp, M. A., Dahlman, K., Grossman, H. T., ... & Beeri, M. S. (2009). The MMSE orientation for time domain is a strong predictor of subsequent cognitive decline in the elderly. International journal of geriatric psychiatry, 24(12), 1429-1437.

Gowans, G., Dye, R., Alm, N., Vaughan, P., Astell, A., & Ellis, M. (2007). Designing the interface between Dementia patients,caregivers and computer-based intervention. *The Design Journal, Volume 10, Issue 1*, 12-23.

Greenwald, B., Cifu, D., Marwitz, J., Enders, L., Brown, A., Englander, J., & Zafonte, R. (2001). Factors Associated with Balance Deficits on Admission to Rehabilitation after Traumatic Brain Injury: A Multicenter Analysis. *Journal of Head Trauma Rehabilitation*, 238–252.

Groppell, S., Soto-Ruiz, K., Flores, B., Dawkins, W., Smith, I., Eagleman, D., & Katz, Y. (2019). A Rapid, Mobile Neurocognitive Screening Test to Aid in Identifying Cognitive Impairment and Dementia (BrainCheck): Cohort Study. *JMIR Aging* , 1-14.

Gustafson, L. (1996). What is dementia? *Neurologica*, https://doi.org/10.1111/j.1600-0404.1996.tb00367.

Hartigan , I. (2006). A comparative review of the Katz ADL and the Barthel Index in assessing the activities of daily living of older people. *International Journal of Older People Nursing*, 204–212. doi:10.1111/j.1748-3743.2007.00074.x .

Herrera-García, J., Rego-García, I., Guillén-Martínez, V., Carrasco-García, M., Valderrama-Martín, C., Vílchez-Carrillo, R., . . . Carnero-Pardo, C. (2019). Discriminative validity of an abbreviated Semantic Verbal Fluency Test. *Dementia & Neuropsychologia*, http://dx.doi.org/10.1590/1980-57642018dn13-020009 .

Hodges, J., & Larner, A. (2017). Addenbrooke's Cognitive Examinations: ACE, ACE-R, ACE-III, ACEapp, and M-ACE. *Cognitive Screening Instruments*, 109-137.

Hodkinson , H. (1972). Evaluation of a mental test score for assessment of mental impairment in the elderly. *Age and Aging*, 233-238.

Hodkinson, H. (1972). Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. *Age Ageing* , 233-8.

Holsinger , T., Deveau , J., Boustani , M., & Williams , J. (2007). Does this patient have dementia?. . *Journal of American Medical Association*, 297: 2391-2404. 10.1001/jama.297.21.2391.

Inoven. (2018). *Cognity*. Retrieved from Home: https://cognity.app

Jitapunkul, S., Pillay, I., & Ebrahim, S. (1991). The Abbreviated Mental Test: Its Use and Validity. *Age and Ageing, 20(5),* , 332–336. doi:10.1093/ageing/20.5.332 .

Johnson, L., Cushing, B., Rohlfing, G., Edwards, M., Davenloo, H., D'Agostino, D., . . . O'Bryant, S. (2014). The Hachinski Ischemic Scale and cognition: the influence of ethnicity. *Age and Ageing, Volume 43, Issue 3*, 364–369.

Jorm , A., Scott , R., & Jacomb, P. (1989). Assessment of cognitive decline in dementia by informant questionnaire. *International Journal of Geriatric Psychiatry*, https://doi.org/10.1002/gps.930040109.

Julayanont, P., Tangwongchai, S., Hemrungrojn, S., Tunvirachaisakul, C., Phanthumchinda, K., Hongsawat, J., . . . Nasreddine, Z. (2015). The Montreal Cognitive Assessment—Basic: A Screening Tool for Mild Cognitive Impairment in Illiterate and Low-Educated Elderly Adults. *Journal of the American Geriatrics Society, 63(12)*, 2550–2554.

Kansagara , D., & Freeman , M. (2010). *A Systematic Evidence Review of the Signs and Symptoms of Dementia and Brief Cognitive Tests Available in VA [Internet].* Washington (DC): Department of Veterans Affairs (US).

Kivipelto, M., Ngandu, T., Laatikainen, T., Winblad, B., Soininen, H., & Tuomilehto, J. (2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, populationbased study . *The Lancet Neurology, 5(9),* , 735–74.

Kivipelto, M., Ngandu, T., Laatikainen, T., Winblad, B., Soininen, H., & Tuomilehto, J. (2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *The Lancent Neurology*, 735-741.

Kohn , L., Corrigan, J., & Donaldson , M. (2000). To err is human: building a safer health system. *Annales francaises d'anesthesie et de reanimation* , 453–454.

Lanza , C., Knörzer , O., Weber , M., & Riepe , M. (2014). Autonomous spatial orientation in patients with mild to moderate Alzheimer's disease by using mobile assistive devices: a pilot study. *Journal of Alzheimers Disorders* , 879-884.

Lei , P., & Koehly , L. (2003). Linear discriminant analysis versus logistic regression: a comparison of classification errors in the two-group case. *The Journal of Experimental Education* , 25-49. .

Leng, F., Yeo , D., George , S., & Barr , C. (2014). Comparison of iPad applications with traditional activities using person-centred care approach: impact on well-being for persons with dementia. *Dementia, London, England*, 265-273.

Maxim, L., Niebo, R., & Utel, M. (2014). Screening tests: a review with examples. *inhalation Toxicology*, 811–828.

Mayo, A. (2016). Use of the Functional Activities Questionnaire. *Best Practices in Nursing Care to Older Adults with dementia*, 323-329.

McCab, D. (2019). *Katz Index of Independence in Activities of Daily Living (ADL).* New York : New York University Rory Meyers College of Nursing.

Melara,, R., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110(3),* , 422–471. https://doi.org/10.1037/0033-295X.110.3.422.

Mioshi, E., Dawson, K., Mitchell, J., Arnold , R., & Hodges, J. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 1078–1085.

Mitchell , A., Bird , V., Rizzo, M., & Meader , N. (2011). Which version of the geriatric depression scale is most useful in medical settings and nursing homes? Diagnostic validity meta-analysis. *American Journal of Geriatric Psycology*, 1066-1077.

Monte, V., Geffen , G., May, C., & McFarland, K. (2010). Improved sensitivity of the rapid screen of mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 28-37.

Moon, H., Badana, A., Hwang, S.-Y., Sears, J., & Haley, W. (2019). Dementia Prevalence in Older Adults: Variation by Race/Ethnicity and Immigrant Status. *The American Journal of Geriatric Psychiatry*, 241-250.

Moroney , J., Bagiella , E., Desmond , D., Hachinski , V., Mölsä , P., Gustafson , L., . . . Tatemichi , T. (1997). Meta-analysis of the Hachinski Ischemic Score in pathologically verified dementias. *Neurology*, 1096-105.

C. F. Moreno-Garc´ıa, E. Elyan, and C. Jayne, (2018) New trends on digitisation of complex engineering drawings," Neural Computing and Applications, pp. 1–18, 2018

Nasreddine , Z., Phillips, N., Bédirian, V., Charbonneau , S., Whitehead , V., Collin , I., . . . Chertkow , H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, 695-699.

Nasreddine, Z. (2019). *MoCA Test*. Retrieved from MoCA Montreal Cognitive Assesment: https://www.mocatest.org/

National Collaborating Centre for Mental Health . (2007). *Dementia: A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care.* Leicester (UK): British Psychological Society.

Newman, C., Bevin, A., Zajicek, J., Hodges, J., Vuillermoz, E., Dickenson, J., . . . Noad, R. (2018). Improving the quality of cognitive screening assessments: ACEmobile, an iPad-based version of the

Addenbrooke's Cognitive Examination-III. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*, 182–187.

Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., . . . Kivipelto, M. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *The Lancent*, 2255-2263.

Nirjon, S., Emi, I. A., Sayeed Mondol, M. A., Salekin, A., & Stankovic, J. (2014). MOBI-COG: A mobile application for instant screening of dementia using the mini-cog test of dementia using the mini-cog test. *5th Conference on Wireless Health.* Bethesda, United States: Association for Computing Machinery, Inc.

Noone, P. (2015). Addenbrooke's Cognitive Examination-III. *Journal of Occupational Medicine*, 418-420.

Okada, S., Inoue, K., Imai, T., & Noguchi, M. (2019). Dementia Scale Classification Based on Ubiquitous Daily Activity and Interaction Sensing. *8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1-7.

Ouimette, P., & Klein , D. (1995). est-Retest Stability, Mood-State Dependence, and Informant-Subject Concordance of the SCID-AXIS II Questionnaire in a Nonclinical Sample. *Journal of Personality Disorders*, 105-111.

Panegyres, P., Berry, R., & Burchell, J. (2016). Early dementia screening. *Diagnostics (Basel)*, doi: 10.3390/diagnostics6010006.

Pang, G.-H., & Kwong, E. (2015). Considerations and Design on Apps for Elderly with Mild-to-moderate Dementia. *International Conference on Information Networking (ICOIN)*, doi:10.1109/icoin.2015.7057910 .

Pazzani , M., & Kibler, D. (1992). The utility of knowledge in inductive learning. *Machine Learning, (9)*, 57-94.

Peters, M. E., & Rabins, P. V. (2017). Cognitive Impairment. In Johns Hopkins Psychiatry Guide Retrieved March 18, 2020, from https://www.hopkinsguides.com/hopkins/view/Johns_Hopkins_Psychiatry_Guide/787027/all/Cognitive_Impairment

Pitarque , A., Roy , J., & Ruiz, J. (1998). Redes neurales vs modelos estadísticos: Simulaciones sobre tareas de predicción y clasificación. *Psicológica*, 387-400.

Poon , T., Chan , A., Zee , B., Ho , S., Mok, T., Leung , T., & Johnson , P. (2001). Application of classification tree and neural network algorithms to the identification of serological liver marker profiles for the diagnosis of hepatocellular carcinoma. *Oncology* , 275-283.

Quinlan, J. (1993.). *C~.5: Programs for Machine Learning. .* Los Altos, Cahfornia, : Morgan Kaufmann.

Robinson , P., & Lowe, J. (2015). Literature reviews vs systematic reviews. *Australian and New Zealand journal of public health*, https://doi.org/10.1111/1753-6405.12393.

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nature Reviews Neurology, 10(11)*, 634–642. doi:10.1038/nrneurol.2014.181 .

Sammut, C., & Webb, G. (2017). *Encyclopedia of Machine Learning and Data Mining.* Springer Publishing Company.

Scharre, D., Chang, S., Nagaraja, H., Vrettos, N., & Bornstein, R. (2017). Digitally translated Self-Administered Gerocognitive Examination (eSAGE): relationship with its validated paper version, neuropsychological evaluations, and clinical assessments. *Alzheimer's Research & Therapy*.

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: the DSM-5 approach. Nature Reviews Neurology, 10(11), 634.

Scharre, D., Chang, S.-I., Murden, R., Lamb, J., Beversdorf, D., Kataki, M., . . . Bornstein, R. (2010). Self-administered Gerocognitive Examination (SAGE): A Brief Cognitive Assessment Instrument for Mild Cognitive Impairment (MCI) and Early Dementia. *Journal of Alzheimer Disease and Associated Disorders*, 64-71.

Shankle, W., Mani, S., Pazzani, M., & Smyth, P. (2005). Detecting very early stages of dementia from normal aging with Machine Learning methods. *Artificial Intelligence in Medicine*, 71-85.

Sindi, S., Calov, E., Fokkens, J., Ngandu, T., Soininen, H., Tuomilehto, J., & Kivipelto, M. (2015). The CAIDE Dementia Risk Score App: The development of an evidence-based mobile application to predict the risk of dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 328-333.

Snyderman, D., & Rovner, B. (2009). Mental Status Examination in Primary Care: A Review. *American Family Physician*, 809-814.

So , A., Hooshyar , D., Park , K., & Lim , H. (2017). Early Diagnosis of Dementia from Clinical Data by Machine Learning Techniques . *Journal of Applied Science*, 1-17.

Statista GmbH. (2018, 10 01). *Statista* . Retrieved from Global mobile OS market share 2009-2018, by quarter: https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/

Tappen, R., Rosselli, M., & Engstrom, G. (2009). Evaluation of the Functional Activities Questionnaire (FAQ) in Cognitive Screening Across Four American Ethnic Groups. *Journal The Clinical Neuropsychologist* , 646-661.

Tariq, S., Tumosa, N., Chibnall, J., Perry, M., & Morley, , J. (2006). Comparison of the Saint Louis University mental status examination and the mini-mental state examination for detecting dementia and mild neurocognitive disorder—a pilot study. *The American journal of geriatric psychiatry*, 14(11), 900-910.

Teng, E., Becker, B., Woo, E., Knopman, D., Cummings, J., & Lu, P. (2011). Utility of the Functional Activities Questionnaire for Distinguishing Mild Cognitive Impairment from Very Mild Alzheimer's Disease. *Alzheimer Disease & Associated Disorders*, 348–353.

Thabtah F., Peebles D., Retzler J., & Hathurusingha C. (2020) A review of dementia screening tools based on Mobile application. Health and Technology (2020) 10:1011–1022

Tombaugh,, T., McDowell, I., Kristjansson, , B., & Huble, A. (1996). Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data. *Journal of Psychological Assessment*, 48-59.

Torre, J. (2004). Is Alzheimer's disease a neurodegenerative or a vascular disorder? Data, dogma, and dialectics. *The Lancet Neurology, 3(3),*, 184–190. doi:doi:10.1016/s1474-4422(04)00683-0

Tsoi, K., Chan, J., Hirai, H., Wong, S., & Kwok, T. (2015). Cognitive Tests to Detect Dementia A Systematic Review and Meta-analysis. *JAMA Internal Medicine*, 1450-1458.

Waszynsk, C. (2012). *The Confusion Assessment Method (CAM).* New York : New York University College of Nursing.

World Health Organization and Alzheimer's Disease . (2012). *Dementia: a public health priority.* World Health Organization.

Yamagata , C., & Kowto , M. (2013). Mobile app development and usability research to help dementia and Alzheimer patients. *Systems, Applications and Technology Conference (LISAT).* Long Island, NY.

Yang, S., Flores, B., Magal, R., Harris,, K., Gross, J., Ewbank, A., . . . Eagleman, D. (2017). Diagnostic accuracy of tablet-based software for the detection of concussion. *PLoS ONE 12*, https://doi.org/10.1371/journal.pone.0179352.

Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, M., & Leirer, V. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 37-49.

Yoganathan, S. (2013, 12 27). *Review of the iPhone app for Delirium and Dementia.* Retrieved from BGS Blog: https://britishgeriatricssociety.wordpress.com/2013/12/27/review-of-the-iphone-app-for-delirium-and-dementia/

Summary Points

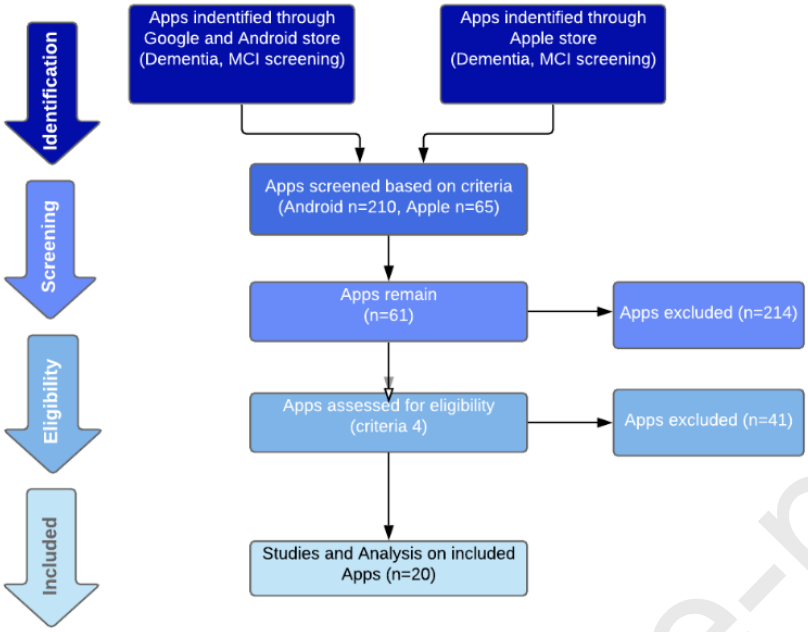
- Multiple dementia diagnosis and screening methods are available on the mobile application stores
- No systematic review for dementia mobile screening apps is available
- New model for mapping dementia screening methods to cognitive areas is proposed

- In depth discussion based on new criteria co compare the available dementia tools is proposed
- Methodology was based on new inclusion and exclusion criteria
- The role of DSM-5 neurocognitive areas was considered in the results analysis
- We consider cognitive tests performance and the role of intelligent methods in the results analysis
- 275 dementia screening tools were considered

On the behalf of the authors I declare that there is no conflict of interest.

Dr Fadi Thabtah

Graphical Abstract



Highlights

- New systematic reviews based on recent criteria for dementia screening methods
- New mapping between dementia screening apps and neurodegenerative cognitive areas
- Performance analysis based on different measures including DSM-5 criteria and comprehensiveness
- 275 mobile screening apps for dementia screening were critically analysed using clear inclusion and exclusion criteria
- Useful mapping to cognitive areas in order to pinpoint to the appropriate app to be used by clinicians in a clinical setting