Evaluating the test-retest reliability of a computerised assessment that measures cognitive ability to drive

Functional outcome assessment following treatment for brain tumour

Belinda Johnston

MSc; BAppSc (Occupational Therapy)

MACQUARIE University

A thesis submitted in fulfilment of the requirements for the degree of

Master of Research

Macquarie University

Faculty of Medicine and Health Sciences

Department of Clinical Medicine

25 April 2018

Table of Contents

Table of Figures

List of Tables

Abstract

INTRODUCTION

This study aimed to evaluate the test-retest reliability of a computerised assessment (DriveSafe DriveAware) that measures cognitive ability to drive. This was with a view to support future research to assess the suitability of DriveSafe DriveAware for use as a repeated measure of high-level cognition that can be quickly administered to brain tumour patients in a clinic or hospital setting.

METHOD

DriveSafe DriveAware iPad version (DSDA) was administered to 40 healthy adults with a current drivers' licence and then repeated 6 weeks and then 6 months later to determine test-retest reliability of scores or whether results changed due to factors such as a learning effect.

RESULTS

A statistically significant improvement in DriveSafe (DS) score and DS completion time was demonstrated between tests 1 to 2. There was no difference between tests 2 to 3 potentially indicating an initial learning effect. The improvement in DS was greater in the participants over 70 years. DSDA completion time was approximately 10 minutes and the improvement in completion time was greater in participants under 70 years. DSDA classification and DriveAware scores remained stable across all three tests.

CONCLUSION

DSDA is a reliable method for repeated assessment of cognition, albeit with a small learning effect between the 1st and 2nd test (but not subsequent testing) affecting the DS score. This effect manifests more noticeably in participants over 70 years of age and is most likely because of increasing familiarity with the technology and decreasing anxiety over repeated assessments rather than because of improved cognition. DSDA should be trialled as a measure to assess high-level cognition of brain tumour patients and as an indicator of whether more comprehensive neuropsychological testing may be warranted.

Statement of Originality

I certify that this thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research and the preparation of the thesis itself have been appropriately acknowledged. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself. In addition, I certify that this thesis has not been previously submitted as part of requirements for a degree to any other institution.

The research presented in this thesis was approved by the Macquarie University Human Research Ethics Committee (Medical Sciences):

<div align="center">

Human Ethics Approval

Reference No: 5201600915 – 2017

</div>

The researcher has no affiliation with Pearson Australia Group Pty Ltd, publisher of DriveSafe DriveAware, and no funding was received from Pearson to complete this research project.

Belinda Johnston

Student no: 45005982

25 April 2018

Acknowledgements

I am so grateful to the many people that have supported me in completing this research. Firstly, I would like to thank Associate Professor Kerrie McDonald and the Cure Brain Cancer Foundation for providing funds for this project through the Sam Ryan Memorial Award. In particular, I would like to thank Andrew and Pam Rorke who generously donated funds to establish this award in memory of their nephew, Sam Ryan, who died from a brain tumour in 2009, aged 31 years.

My principal supervisor, Associate Professor Andrew Davidson, has been a constant source of encouragement and infectious enthusiasm. I am very appreciative of his time and his wealth of knowledge that he has so generously provided, despite his incredibly busy schedule. I cannot thank him enough for taking a chance and giving me the opportunity to dip my toes back into the world of research and academia.

My co-supervisor, Professor Marcus Stoodley, has been an invaluable resource and provided a great deal of support and oversight to my principal supervisor. Dr Maurizio Manuguerra generously gave his time to assist me with the statistical analysis.

Thank you to the Higher Degree Research; Faculty of Medicine and Health Sciences; and Macquarie Neurosurgery staff, who were always ready to help out whenever I had an issue. In particular, I would like to thank Dr Jennifer Rowland, Mary Simons, Viviana Bong, Collette Tosen, Kate Morkel, and Lesley Morris.

A huge thank you goes to all my friends who supported me, listened, volunteered as participants, read through drafts and shared a glass of wine at crucial times. You know who you are! And of course, to Joan O'Donnell who encouraged me to take this on and then made sure I kept going.

Most importantly, I am extremely grateful to my immediate family: Ellie and Aidan – who made cups of tea as required and even borrowed books from their respective universities for me. But mainly to my endlessly supportive husband, David, who not only encouraged me to take up this opportunity but was my main champion along the journey. Thank you.

1. Introduction

DriveSafe DriveAware (DSDA) is an assessment developed to predict if a person with cognitive impairment is safe to drive (1). Future research seeks to administer DSDA to assess brain tumour patients at various points along their treatment pathway, with the goal of assessing whether DSDA assists in detecting changes in high-level cognition. Therefore, the test-retest reliability of the assessment must first be established.

1.1 Brain tumour and the impact on driving

Brain tumour is a devastating diagnosis. Both the disease and the treatment may have a significant impact on cognition and quality of life. Survival rates, especially for glioblastoma, remain low but with commitment to ongoing research from governments and the medical community, progress is anticipated (2). As survival rates improve, consideration must be given to outcome measures that look beyond survival measures and focus on functional measures and quality of life (3, 4). Driving is an important functional task but patients with brain tumours are often excluded from this key activity of daily living. Driving restriction contributes to reduced independence and quality of life (5).

1.2 Brain tumours

1.2.1 General

Brain tumours are a heterogeneous group of neoplasms affecting the central nervous system and include primary brain tumours and metastatic tumours. Primary brain tumours arise from the brain and associated tissues and are broadly categorised as malignant and non-malignant (benign) tumours. The most common primary brain tumour is glioblastoma. Metastatic brain tumours arise from multiple types of cancer but most commonly arise from lung cancer, breast cancer, melanoma, colon-colorectal or kidney/renal cancer. Brain tumours occur in all life stages, but the prevalence rises after 45 years of age, peaking at 75-79 years of age (6).

Brain cancer is the sixth highest cause of cancer burden. The Australian Institute for Health and Welfare (AIHW) estimates that in Australia, 2 076 new cases of brain and other central nervous system (CNS) cancer will be diagnosed in 2017, and 1 477 people will die from this

disease (6). It is estimated that in 2018, there will be 23 830 primary malignant brain tumours diagnosed in the USA (7).

Gliomas most commonly occur in the frontal lobe, the temporal lobe, and the parietal lobe respectively (6). Functions of these lobes include executive functions, memory, and integration of sensory input - crucial elements for cognition (8, 9). Glioblastoma (GBM) is the most common and most severe form of malignant glioma. It has exceedingly detrimental effects on a patient's health, independence, and quality of life (10).

In Victoria, Queensland and Western Australia, 1 029 cases of benign tumour of the brain and other CNS were diagnosed in 2013 – the majority (66%) being in the meninges. In 2015, 279 deaths were reported in Australia as a result of benign tumours of the brain (6). Benign tumours generally have distinct borders, are slow growing, and rarely spread but they can be fatal if located in a critical location in the brain. In the USA, it is estimated that 55 150 benign tumours will be diagnosed in 2018 (7).

The numbers of malignant and non-malignant tumours may appear low in comparison to other cancers and health conditions, but the health burden to patients, their families, and the wider community in terms of treatment and years of life lost is considerable (6). On 28 October 2017, the Australian Federal Government announced the *Australian Brain Cancer Mission* in association with the Cure Brain Cancer Foundation, and the Minderoo Foundation's Eliminate Cancer Initiative. The funding and collaboration aims to support research and action to double survival rates in 10 years and improve quality of life (11) and is an indication of the Australian Government's commitment to reducing the incidence and impact of brain cancer.

1.2.2 Treatment

Initial treatment for malignant gliomas typically involves cortico-steroids, anticonvulsive therapy (if required), and surgery (for biopsy and possible resection). Tumour histology guides decision-making regarding follow-up chemotherapy or radiotherapy (12). Standard therapy for patients with glioblastoma includes surgery, radiation therapy and chemotherapy (13).

Treatment of metastatic brain cancer follows cancer specific algorithms and must be managed in conjunction with treatment for the primary cancer (14). Treatment options usually include surgery, radiotherapy, chemotherapy or a combination of these modalities.

New advances in tumour biology have revealed opportunities for immunotherapy to treat GBM (15). There are significant challenges, although trials are exploring vaccines, checkpoint inhibitors, adoptive T cells, and combinations of therapies. It is anticipated that various approaches in combination with targeted immunotherapies will be required in new treatments focused on GBM and other brain tumours (15).

Benign tumours are treated either surgically, or if asymptomatic, monitored for growth, symptom development and effects on function. If removed, benign tumours may regrow but unlike some malignant tumours, do not metastasise to other parts of the body.

1.2.3 Survival

Benign tumours have a relative 5-year survival rate of 95 to 100%, only reducing at age 65+. In contrast, survival rates for malignant tumours are low and 25% of those diagnosed will not survive 5 years. This rate has remained steady for the last 30 years despite increased survival rates in other forms of cancer (6). The median survival for patients with GBM is 12 to 15 months (13). However, as treatment regimens develop, improvements in future survival rates are expected (2).

As survival rates improve, future function and quality of life for a brain tumour survivor is critical to consider in their treatment planning. The disease process and the subsequent treatment often lead to cognitive deficits that negatively affect patient function and quality of life. However, it is essential that longer-term survival is accompanied by good quality of life (3, 4).

1.3 Brain tumour and impact on cognition and quality of life

Cognitive dysfunction significantly impacts activities of daily living and quality of life (16). Cognitive dysfunction occurs in brain tumour patients due to both tumour growth and subsequent treatments such as radiotherapy (17). Stroke patients may have cognitive deficits concentrated to a specific part of the brain. In contrast, patients with a glioma have

deficits that are more general. This is possibly due to diffuse growth of the tumour and cells infiltrating healthy brain matter. Cognitive deficits are greater when the glioma is in the dominant hemisphere than when the lesion is in the non-dominant hemisphere (17). The deficits experienced by glioma patients include working memory, cognitive control and flexibility, cognitive processing speed, visual searching, planning and foresight, and general attention, independent of other variables such as age, gender, and education (18).

Deterioration in cognition may be an early sign of tumour recurrence after treatment and cognitive testing may alert the clinician before radiological evidence (19). Cognitive assessment may also be a valuable predictor of survival in some forms of brain tumour (20, 21). Clinicians should consider assessment of cognition as a priority within the range of outcome measures available to them.

1.4 Outcome assessments

Overall Survival (OS) or Progression Free Survival (PFS) are the long-established measures of the success of neurosurgical intervention for brain tumour. OS is considered the "gold standard primary end point to evaluate the outcome of any procedure that is assessed in oncologic clinical trials" (22). However, people are living longer after diagnosis and treatment for other forms of cancer, so the focus is turning towards future planning for the survivor (23) and the functional outcomes post treatment. The term "cancer survivor" is used in various contexts and often refers to someone who has finished their primary treatment. Adding other measures such health-related quality of life (HRQoL) or instrumental activity of daily living (IADL) measures gives a more complete picture of outcome in the context of survivorship (24). As brain tumour treatments improve survival rates, Clinical Outcome Assessment (COA) should be included routinely with OS, PFS and radiological and biomarker assessments (25, 26). A COA is one that is influenced by human judgement and choices. It relies on interpretation and reporting from a patient, a clinician, or an observer (27).

COAs include:

- patient-reported outcome measures;
- clinician-reported outcome measures;
- observer-reported outcome measures; and
- performance outcome measures.

1.4.1 Patient-reported outcome (PRO) measures

PROs are regularly used worldwide with cancer patients (28). The Short-form (36) health survey (SF-36) (29), the European Organisation for Research and Treatment of Cancer (EORTC) QLC-C30 (30), and the Functional Assessment of Cancer Therapy - General (FACT-G) (31) scale are the most routinely used. The BN-20 survey combines with the QLQ-C30 to specifically assess outcome in brain cancer patients, as does the Functional Assessment of Cancer Therapy-Brain (FACT-Br). These surveys assess quality of life, symptoms, health perceptions, and physical and cognitive functional ability. Cognitive dysfunction may affect data accuracy due to patient altered perception of function versus the reality of function. For example, a patient with cognitive impairment may report that they are independent for self-care, but a carer may indicate that the patient requires significant assistance. Reviews show some success at using proxies to complete these QoL assessments, however proxy use should only be considered if there is no prospect of the patient completing the survey independently (4).

1.4.2 Clinician-reported outcome (ClinRO) measures

A ClinRO is an assessment administered by a clinician. The Neurologic Assessment in Neuro-Oncology (NANO) working group recommends a set of assessments to standardise neurological evaluation of brain tumour patients. The nine assessed domains are: gait, strength, upper extremity ataxia, sensation, visual fields, facial strength, language, level of consciousness, and behaviour. Each domain receives a score and the cumulative score can be compared with previous assessments. The battery of assessments demonstrates high inter-observer reliability (25). Although each is an objective measure in isolation, and useful when the patient has significant deficits, these domains do not examine ability to perform high-level functional activities.

Clinicians regularly use the Karnofsky Performance Scale (KPS) (32) to assess pre-treatment and post treatment function in tumour patients. The clinician subjectively assesses whether the patient is "able to carry on normal activity" but there are no specific criteria on which to base an assessment of "able". The KPS does not discriminate ability to perform high-level cognitive tasks, such as work participation or driving.

1.4.3 Observer-reported outcome (ObsRO) measures

An ObsRO is information reported by an observer who is not the clinician or other health professional. Medical judgement is not required. An ObsRO is considered an important element of patient engagement as it employs others in the community to contribute to assessment (33). For a brain tumour patient, an ObsRO assessment may be provided by a friend, carer or family member reporting observable behaviours or symptoms such as vomiting, drowsiness, wincing in pain, or the degree of assistance required for self-care activities.

1.4.4 Performance outcome (PerfO) measures

A PerfO is measured by a health professional who directs the patient to perform a task. This may be a physiotherapist conducting a gait test, or a neuropsychologist performing memory or other cognitive testing. PerfOs to assess cognition are vital for patients with brain tumour because cognitive dysfunction is a key symptom of both malignant and non-malignant brain tumour.

The International Cognition and Cancer Task Force recommend a battery of neurocognitive tests (PerfOs) to standardise testing and enable comparison of patient data across studies. These tests are the Hopkins Verbal Learning Test-Revised (HVLT-R), Trail Making Test (TMT), and the Controlled Oral Word Association (COWA) of the Multilingual Aphasia Examination (16, 34). The Montreal Cognitive Assessment (MoCA) and the Mini Mental State Examination (MMSE) are also commonly used in assessment of cognition in brain tumour patients in international studies. The MoCA is superior to the MMSE as a prognostic indicator and has better correlation with quality of life measures (35) although according to Naehrig et al., by 2016, the MOCA had not been widely adopted in Australian oncology practice (36). Robinson and colleagues (37) perhaps explained this when they reported that in patients with brain tumour, the MoCA had poor sensitivity for detecting cognitive impairments and recommended a brief but tailored cognitive assessment by a neuropsychologist. Curiously, none of these psychometric tests relate to a specific IADL (functional) task that requires high-level cognition. Additionally, administering the entire battery of tests to a patient with a brain tumour may be a lengthy and a fatiguing experience for the patient. Gehrke et al. (18) argue that future research should focus on and evaluate the functional challenges of brain tumour survivors and in particular those that result from

"cognitive sequelae". Their conclusion is that "future management of these cases needs to go beyond tumour surveillance" and health providers need to play an active role in optimising function.

1.4.5 Novel approaches to functional outcome assessment

Activity of daily living assessments in current use with brain tumour patients include the Modified Barthel Index (MBI) (38) and the Functional Independence Measure (FIM) (39). While providing valuable information regarding functional capacity, the focus of these assessments is on self-care activities. The FIM also includes cognitive elements such as comprehension, expression, social interaction, problem solving, and memory; however, the MBI and the FIM do not assess capacity to perform IADLs that require higher-level cognitive function. A formal driving assessment should be considered to assess capacity to drive.

The Perceive, Recall, Plan and Perform (PRPP) system of task analysis (40) may be a valuable model in which to consider the complexity of the driving task as the elements required for driving fall into each of the four quadrants described within the framework. The PRPP has been successfully used to assess breast cancer survivors experiencing cognitive problems when returning to the work place (41). However, no literature could be identified that specifically used the PRPP model to analyse driving.

Driving is a form of community mobility recognised by the International Classification of Functioning, Disability and Health (ICF) (42). Driving is a critical area of occupation that occupational therapists should address with each client (43) and it is a key activity for older people as it significantly contributes to independence (44). Indeed, an occupational therapy research survey of older people identified driving (amongst other IADLs) as the most important activity (45). Driving is also a key activity in younger years as it offers independence and is considered a rite of passage into adulthood. Using a functional assessment such as driving ability may improve patient compliance with follow up assessments as the assessment specifically relates to a task the patient is interested in resuming.

No literature could be identified that used driving ability as a measure of cognition following diagnosis and treatment for brain tumour.

1.5 Brain tumour and driving

1.5.1 Requirements for driving

The research literature is limited regarding driving after diagnosis and treatment for brain tumour (46, 47); however, driving has been widely researched in relation to other neurological conditions, especially traumatic brain injury (TBI), stroke, epilepsy, and dementia. Differences arise in the disease course for a patient with a brain tumour, but the potential cognitive deficits are similar enough to assume that an instrument to assess cognition following TBI or stroke may be suitable for assessing cognition with a brain tumour. Whether due to physical or cognitive impairment, licensing authorities generally require authorisation from a medical practitioner in cases where concerns arise around a patient's capacity to drive. Limited guidance or education exists for medical practitioners on the specific evidence on which to base their decision as to whether a patient is fit to drive (48). A neuropsychologist may provide an opinion on cognitive capacity to drive, but in Australia an occupational therapist more commonly provides expert evidence to clinicians on a patient's physical and cognitive capacity to drive.

1.5.2 Existing guidelines for driving and brain tumour

A Cancer Council Australia guide for patients with a brain tumour and their families explains that driving may be restricted on a medical practitioner's recommendation following diagnosis and treatment of a brain tumour. It warns that patients may feel frustrated by the reduced independence due to driving restrictions (49). Clinician surveys demonstrate an inconsistent approach to assessing the capacity to drive safely after brain tumour, and accordingly, many researchers recommend the development of clear guidelines for clinicians (50-52). An Australian guideline provides advice for clinicians in regard to driving and brain tumour, but is more advisory than prescriptive (5, 53). It recommends action based on neurological deficits but does not provide direction on how to identify or measure the deficits.

Driving restrictions are addressed in the recently updated Austroads guidelines (54). The Austroads guidelines state that a person with neurological deficits from a space occupying lesion is unfit to hold an unconditional licence but may hold a conditional licence with medical authorisation. The recommended driving restriction after intracranial surgery is six

months, but this is designated "advisory only" and relies upon the opinion of the neurosurgeon.

The Austroads guidelines' default guidance for private drivers in relation to seizure is a 12-month restriction from driving after the most recent seizure. Many patients who experience seizures, whether due to the disease process, or in response to treatment, will be prescribed anti-epileptic drug therapy and restricted from driving for six months by their clinician (55). This aligns with the Austroads recommendation that a conditional licence can be issued six months after the last seizure with a medical expert recommendation. When seizures are associated with another neurological condition, both should be considered when making a driving recommendation.

The medical practitioner determines whether a patient's cognitive impacts are too severe for safe driving. Apart from referral to a qualified occupational therapist driver assessor, no specific guidance is available for a medical practitioner on which evidence to base the cognitive assessment.

1.5.3 Occupational therapy driving assessments

In Australia, an occupational therapy driving assessment generally involves both an on-road and off-road assessment. The occupational therapist assessor requires post graduate specialist training in driver assessment. However, because driving is a key IADL, Dickerson et al. state that "all occupational therapists, regardless of their area of practice, need to understand and use screening and assessment tools in the areas of cognition, motor skills, vision, and perception to advise the medical team on fitness to drive" (56). The occupational therapy driving assessment can be time consuming, and the expense and effort required may be considered unreasonable if the patient is at an early stage of their treatment, especially if their cognition is likely to deteriorate or improve. A screening tool should be used to identify those patients who would benefit most from an occupational therapy driving assessment to reduce the time and cost burden and provide support information for clinicians in their decision as to whether a patient is safe to drive (47, 57).

1.5.4 Assessments that predict fitness to drive

The occupational therapy on-road assessment remains the gold standard for assessing fitness to drive (56). Conversely, researchers acknowledge that no gold-standard off-road assessment (that is an assessment performed in a clinic rather than in an automobile) exists to accurately predict, in all circumstances, whether a person is safe to drive (58). A benefit of an off-road assessment is that it is likely to be quicker and less resource intensive than an on-road assessment.

Over the last 10 years, there have been multiple research projects and systematic reviews aimed at identifying an off-road assessment or battery of assessments that can accurately predict whether a person has the cognitive capacity to safely drive an automobile (56, 58-80). A similarity of these assessments is their reliance on commonly used psychometric tests, and many of the studies focus on the correlation of performance on the psychometric tests with on-road driving ability (usually an on-road assessment). Tests regularly reported in these batteries include the MMSE, MoCA, and the Trail Making test.

Executive function, processing speed, visuospatial skills, attention, memory, and mental flexibility are identified as predictive of fitness to drive (81). Insight, awareness, and self-regulation also influence cognitive fitness to drive (82-86). A person who has insight that their driving skills have altered and is aware of their errors is more likely to self-regulate behaviour (e.g. drive at quieter periods, avoid night driving, proceed with extra caution at an intersection) and consequently, is safe to drive, or at least safe with the restrictions of a conditional licence.

DriveSafe DriveAware (DSDA) is an off-road assessment specifically designed to assess cognitive fitness to drive and may be an alternative to administering general psychometric tests to predict cognitive function. If a patient has previously held a valid driving licence, it is reasonable to assume they were competent driving an automobile prior to their disease. A driving assessment allows a judicious comparison between current cognitive function and cognitive function prior to their disease.

1.6 DriveSafe DriveAware

DSDA is a two-part cognitive screening tool: DriveSafe (DS) and DriveAware (DA), and was developed by occupational therapists in Australia, for use with patients with suspected cognitive impairment, to predict their likelihood of passing an occupational therapy on-road assessment. The two scores (DS and DA) combine to measure a driver's awareness of the driving environment and their own abilities related to driving (1). It takes approximately 10 to 15 minutes to complete and assesses the elements of executive function, processing speed, visuospatial skills, attention, memory, and mental flexibility. DriveAware assesses awareness and insight. The combined score classifies drivers into *"safe", "unsafe"*, and *"needs further testing"*. The original version of DSDA had both specificity and sensitivity over 90% when compared to an occupational therapy on-road assessment (57, 78).

The classification of "*safe"* predicts that a person is *"likely to pass an occupational therapy on-road assessment"*, and *"unsafe"* predicts that a person is *"likely to fail an occupational therapy on-road assessment"*. *"Needs further testing"* indicates the person should be referred for an occupational therapy driving assessment as they may be cognitively unfit to drive, or there may be a need for training or adaptions to ensure their driving safety. This recommendation reinforces that an occupational therapy on-road assessment is the definitive measure of whether the person is safe to drive.

Combining the DriveSafe and DriveAware scores reaffirms that safe driving is a combination of attention, memory, and visual-spatial ability as well as insight into one's own ability to drive (77). Therefore, it is possible that a patient with a high DriveAware score may still be found *"likely to pass an occupational therapy on-road assessment"* despite a marginal score in the DriveSafe section. But if the assessor, relying on clinical judgement, continues to have doubts regarding cognitive capacity to drive, then the patient is referred for an occupational therapy driving assessment.

Originally developed as a computer-based assessment with an administrator recording answers with pen and paper, DSDA was modified in 2015 for administration with an iPad® (1). The revised DSDA is also known in the research as "DSDA for touch screen", "touch screen DSDA", "DSDA touch screen version" or "DSDA iPad-based version" (2015).

The DriveSafe component of the DSDA for touch screen involves showing participants a series of 10 images of an intersection. Objects (cars, pedestrians, other road users) appear on the intersection. After a few seconds the objects disappear. The participant is required to touch the screen and indicate where the object was, select the type of object, and drag the object in the direction that it was traveling. The score is based on the number of objects that the participant correctly identifies. The maximum score is 84.

DriveAware is a series of questions administered both through onscreen questions and questions posed in an interview with the administrator (a qualified health professional). The answers to the onscreen questions are selected on the screen by the participant. The answers to the interview questions are selected on the screen by the administrator. The maximum score on DriveAware is 17. The two scores are combined to form the DSDA score and recommendation.

The DriveSafe component can be supervised by a clinic staff member, such as a research assistant or practice nurse, who understands the administration procedures outlined in the DSDA manual (1). The DriveAware section of the assessment, which includes a patient interview and clinical judgement, must be administered by a qualified assessor. A qualified assessor is a health professional such a medical practitioner or registered occupational therapist, physiotherapist, psychologist, or speech pathologist. Post graduate specialist driving assessment training is not required. In practice, the research assistant may commence the assessment with the patient and then a qualified assessor can administer the DriveAware interview. Gooden et al. (82) raises concerns that DriveAware relies on patient reporting of recent driving experiences and that the patient may not have driven an automobile for some time. This is a valid concern but the DriveAware questions relate to the iPad task (series of intersections) and the health professional can exercise clinical judgement to influence the DriveAware score during the interview.

A prospective study administered DSDA touch screen version to a convenience sample of 134 older (65+) and cognitively impaired (18 year +) drivers in Australia and New Zealand (87). Participants were administered DSDA and the Mini-Mental State examination as part of an off-road assessment. A standardised on-road assessment was used as the criterion measure for this study. 112 participants were used for the predictive validity study. The

optimal cut off sores on the DriveSafe subset were 57 and 72 and were 10 and 13 for the DriveAware subtest. Cheal and Kuang (1) report that, "the test identified unsafe drivers at the low cut-off score with Sensitivity of 91% (95% CI: 84 to 96) and 89% (95% CI: 82 to 97) for DriveSafe and DriveAware, respectively. The test identified safe drivers at the upper cut-off score with Specificity of 94% (95% CI: 87 to 99) and 91% (95% CI: 84 to 99) for DriveSafe and DriveAware, respectively."

The study's predictive validity results demonstrated that DSDA has a Specificity of 86% and Sensitivity of 91%, a Positive Predictive Value of 83%, Negative Predictive Value of 92% and an overall accuracy of classification of 88% (that is classification as "safe", unsafe "or "requires further testing") (1). To date there are no published critical appraisals of these results.

1.6.1 Test-retest reliability and learning effect

Test-retest reliability data has not been published for the DSDA iPad version. An assessment of the test-retest reliability of the DS score on an earlier (non-iPad) version of DSDA found good intraclass correlation of 0.67 (p < 0.01) (88) but comparison with the iPad version is problematic due to the differences in scoring and administration. There was no evidence in the literature as to whether there is a learning effect when the DSDA iPad version is repeated and whether unfamiliarity with touchscreen technology may affect repeated scores.

A population study undertaken in Brazil demonstrated an improvement in a battery of cognitive tests performed in 160 adults (with a mean age of 52 years), suggesting that better performance in tests of cognition may occur as the result of a learning effect (89). A Melbourne study assessed the performance of 113 neurologically normal adults (mean age 64 ± 8 years, range 46–82 years) on repeated attempts of a cognitive test and found a practice effect between the 1st and 2nd administration of the test battery, with smaller, nonsignificant improvements observed between the 2nd, 3rd, and 4th administrations (90).

The test-retest reliability of DSDA must first be established to support future research that seeks to administer DSDA to assess brain tumour patients at various points along their

treatment pathway, with the goal detecting changes in high-level cognition while also predicting whether they are likely to pass an on-road occupational therapy assessment.

1.7 Tablet technology and the elderly

It is often assumed that older people (aged 70+ years) will struggle with using technology, unlike the so-called digital natives (91) that have been born into a world that uses technology, and younger digital immigrants (91) who have learnt about technology after their youth but have adapted well using it in their everyday home and working life.

As populations of developed nations age, older people are the fastest growing demographic of computer technology users (92). Despite this, there are concerns about how older people will adapt to tablet technology. Performing tasks on a tablet that they are already confident to perform on a computer may be difficult for some older people due to the different technical interface required, such as the keyboard on a tablet seemly "appearing and disappearing" (93). However, a study assessing the usability of tablets for people with early stage dementia indicated that the more intuitive touchscreen interface simplifies interactions due to the absence of an external controller such as a mouse (94) and 48% of the study participants reported that the tablet technology was 'moderately' or 'extremely' intuitive.

1.8 Summary

Brain tumour continues to be a devastating diagnosis to both individuals and the community. However, as treatments develop and improve, and survival rates increase, focus should be on functional outcome and quality of life. Cognitive dysfunction affects ability to perform IADL and reduces quality of life. Cognitive assessment may be predictive of functional outcome and survival. Assessment of a significant functional task that requires high level cognition, such as driving, should be considered as a screening tool before lengthier neuro-psychological testing. Such an assessment may also detect deterioration or tumour regrowth before radiological evidence. DSDA is a succinct off-road driving assessment of higher cognition. It may be a novel tool to detect cognitive changes (whether improvement or deterioration) in patients with brain tumour. It may also provide objective evidence regarding a patient's competency to drive an automobile.

2. Aims of the Study

This thesis focused on assessing the test-retest reliability of DriveSafe DriveAware iPad version (DSDA). Research indicates that DSDA, when used as a one-off test, is a reliable and valid functional measure of a driver's awareness of the driving environment and cognitive ability to drive. Its predictive value (that is – ability to ability to predict whether a driver with cognitive impairment is cognitively safe to drive or whether they require an on-road occupational therapy assessment) is high (Positive Predictive Value = 83%; Negative Predictive Value = 92%) (1). DSDA's Specificity is 86%, Sensitivity is 91% and overall Accuracy of Classification is 88% (1, 87) when compared with (the gold-standard) occupational therapy on-road assessment. No current research exists however to indicate whether DSDA iPad version is reliable when used as a repeated measure.

The primary aim of the study was to assess whether DSDA is a reliable measure when repeated on healthy subjects. The secondary aims were to assess whether different age groups performed differently on repeated testing and whether gender affects DSDA outcomes.

This study was conducted as a part of a larger Macquarie University project that aims to identify suitable outcome measures for brain tumour patients and aims to assess DSDA's possible utility to detect changes in the cognition of brain tumour patients prior to surgery and then at repeated follow-up clinic appointments. There is a prospect that DSDA may be trialled as a measure of change in high level cognition for patients undergoing treatment for brain tumour, in addition to predicting whether they are likely to pass an occupational therapy on-road assessment.

2.1   Research Question

This study examined the following research question:

Is DriveSafe DriveAware iPad version a reliable measure when repeated on healthy subjects?

2.2   Research Hypotheses

This study examined the following research hypotheses:

That the DSDA results of healthy people do not differ over repeated assessments.

That there is no difference in results between genders over repeated assessments of the DSDA.

That there is no difference in results between age groups over repeated assessments with DSDA

The four DSDA components assessed were DSDA classification; DriveSafe scores; DriveAware scores and DriveSafe completion time (elapsed time).

3. Methodology/Materials

The study used a quantitative methodology to assess the test-retest reliability of DSDA iPad version by administering the same test on three occasions to the same 40 healthy participants and the data was analysed to determine whether scores remained the same or improved, despite no self-reported changes to cognition. The STROBE guidelines for reporting observational studies were used in the preparation of the methodology (95) and the study was approved by the Macquarie University Human Research Ethics Committee Human Ethics Approval (Reference No: 5201600915 – 2017).

3.1 Participants

Forty volunteers (17 male; 23 female) were recruited for the study from the south-eastern region of Australia (predominantly the Australian Capital Territory and New South Wales). The median age was 58 (range 20-91 years) and the completed education level ranged from year 10 equivalent to University Master's Degree.

The participants all met the following inclusion criteria: held (as a minimum) a current class C drivers' licence (an Australian state-based licence that permits driving of an automobile or light truck up to 4.5 tonne and that seats up to 12 adults, including the driver); aged 20 years or over at the time of first assessment (in the majority of Australian states drivers cannot hold a class C licence until 20 years of age or older); reported that they drove an automobile regularly; demonstrated English proficiency; and self-reported as being in good health with no cognitive deficits. Participants were excluded if they reported they were not available for all three assessments; did not self-report as regular and confident drivers; or if they self-reported as suffering from a current disease or used medication that may have affected cognition.

The participants were a convenience sample drawn from willing volunteers approached by the researcher, and their friends, family and work colleagues. No payments or inducements were offered for participation.

The sample size was determined based on the intraclass correlation coefficient (ICC) reported in a study assessing test-retest reliability for the original version (not touchscreen)

of DSDA (88). This previous study of test-retest reliability had an ICC of 0.67 (p < 0.01) (88). Bujang and Baharum (96) estimate that the sample size required for ICC of 0.6; with power = 90%; alpha = 0.05; and 3 observations per participant, is N = 10. A sample size of 40 was planned for this study to allow for redundancy and to allow the sample to be divided into three age groups for the analysis. More participants aged 50+ years were recruited than under 50 years as there is a higher risk of development of brain tumour in those over 50 years of age - the median age for the incidence of all brain tumour is 59 years of age (7). Participants were deliberately approached to ensure the sample adequately represented a broad age range (from 20 to 91 years). More volunteers than required were available in the 50 to 59-year-old range but were excluded due to high numbers in this age range and because not all would be available for all three assessments.

It was difficult to recruit willing volunteers aged over 70 years and particularly over 80 years who had no cognitive deficits, were in good health and were still driving regularly, despite approaching several potential volunteers. A couple of prospective volunteers were excluded from the study due to ill health or regular medication that had the potential to affect their cognition. Ten participants represented this age group in the study. Prior to the second and third assessments the participants were asked to re-confirm their previous self-report as being in good health with no cognitive deficits and no current medication that may affect their cognition. An 80-year-old male was excluded from the study prior to the third assessment due to ill health but the results from his first two tests (when he was in good health) were included in the analysis between test one and test two.

3.2 Materials and Assessment Tools

All assessments were completed on the same iPad Air 2 with a 9.7-inch screen. The iPad was in a case that allowed it to be set up at a 20-degree angle in front of the user. The iPad was loaded with the DriveSafe DriveAware (DSDA) program (iPad version 2015), purchased from Pearson Clinical. The DSDA program produced an individual report for each subject on the completion of each test.

3.3 Procedure

The researcher (a registered occupational therapist) administered all assessments in the research project. Prior to the first assessment, the purpose of the research project and the subject's rights and obligations were explained by the researcher. Each subject signed a consent form approved by the Macquarie University Human Ethics Research Committee. The form explained the purpose of the research study and outlined the participant's agreement to complete the DSDA assessment on multiple occasions with results collated and reported for research purposes. The researcher assured confidentiality and informed the participants that withdrawal from the study was permitted at any time. The researcher informed the participants that they would be referred to their general practitioner if the assessment predicted they *"required further testing"* or were *"unlikely to pass an occupational therapy on-road assessment"*. Demographic information on the participants was entered on the iPad into the DSDA application.

The assessment procedure followed the General Administration Guidelines in the DriveSafe DriveAware (DSDA) Administration Manual (1). The subject was seated at a table in a quiet room without distractions. The iPad was placed on the table, directly in front of the subject (in their midline) at a 20-degree angle and the researcher sat beside the subject on their non-dominant side. The researcher provided minimal guidance during the practice phase of the test, if required. After the assessment commenced there was no assistance provided. In response to requests for assistance, the researcher stated, "sorry, I cannot provide assistance during the test. Just do your best," as recommended in the DSDA Administration Manual. The DSDA Administration Manual outlined an Administrator Assisted Method for use when the person cannot self-administer the test. This study did not use the Administrator Assisted Method at any time. Headphone and stylus use was outlined in the manual, but they were not used in this study. In some instances, if the subject appeared comfortable completing the self-administered section, the researcher left the room and returned when the DSDA audio instructions asked the subject to "please hand the iPad to your administrator". A clinician interview followed the self-administered DSDA assessment on the iPad and the researcher made the clinical assessment based on the response to the interview questions.

Repeat assessments followed an identical procedure, approximately 6 weeks (42 days) and then approximately 6 months (183 days) after the first assessment. The median time from test one to test two was 43 days (range 40 – 57 days) (*Figure 1*). The median time from test one to test three was 183 days (range 171 – 198) (*Figure 2*). Fifty-seven-point five percent (57.5%) of participants completed their second test either on the 42-day mark or one day either side of that date. Forty-four percent (44%) of participants completed their third test on the 183-day mark or one day either side of that date. The variability in time frame of the assessments aligned with the variability expected with patient follow up appointments.

Time discrepancy from initial test to 6 week follow up test
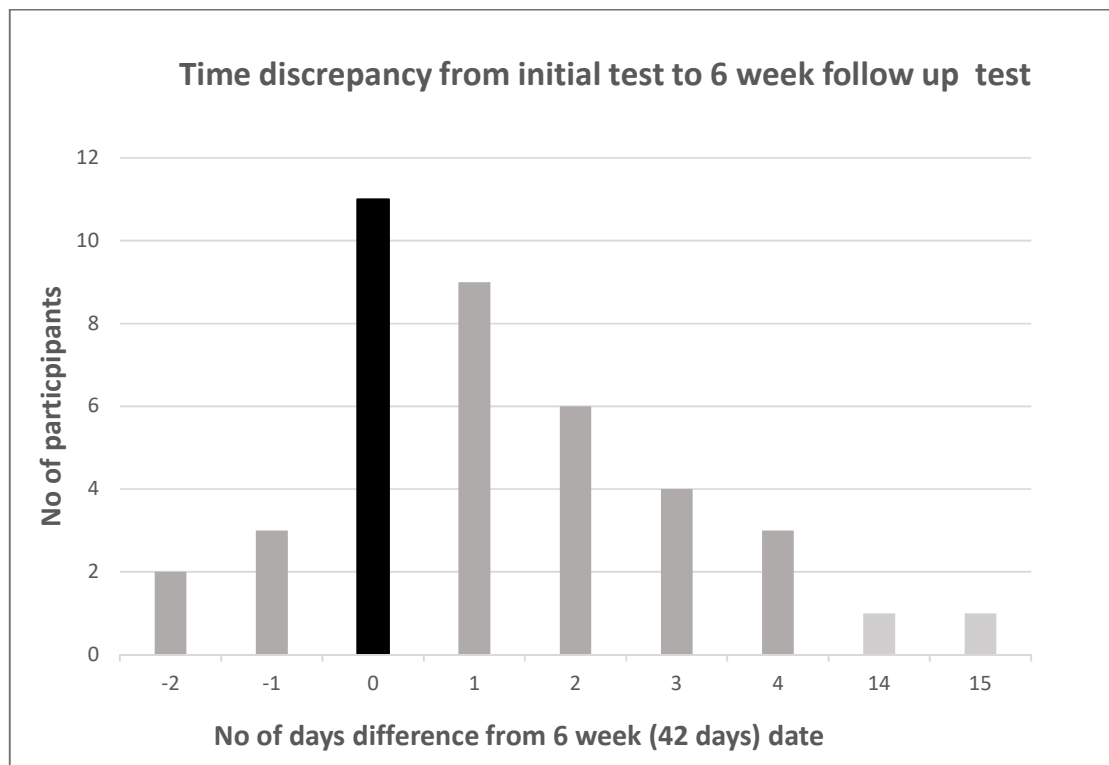
Figure 1. No of days from nominated test 2 date at 42 days (marked as zero) and the actual date of test 2

Participants were blinded to their DriveSafe and DriveAware test scores, and their DriveSafe completion time score during the research project. However, for ethical reasons, on completion of the assessment, the participants were told whether their DSDA classification result was *"likely to pass an occupational therapy on-road assessment"*, *"likely to fail an occupational therapy on-road assessment"*, or *"requires further testing"*.

Time discrepancy from initial test to 6 month follow up test
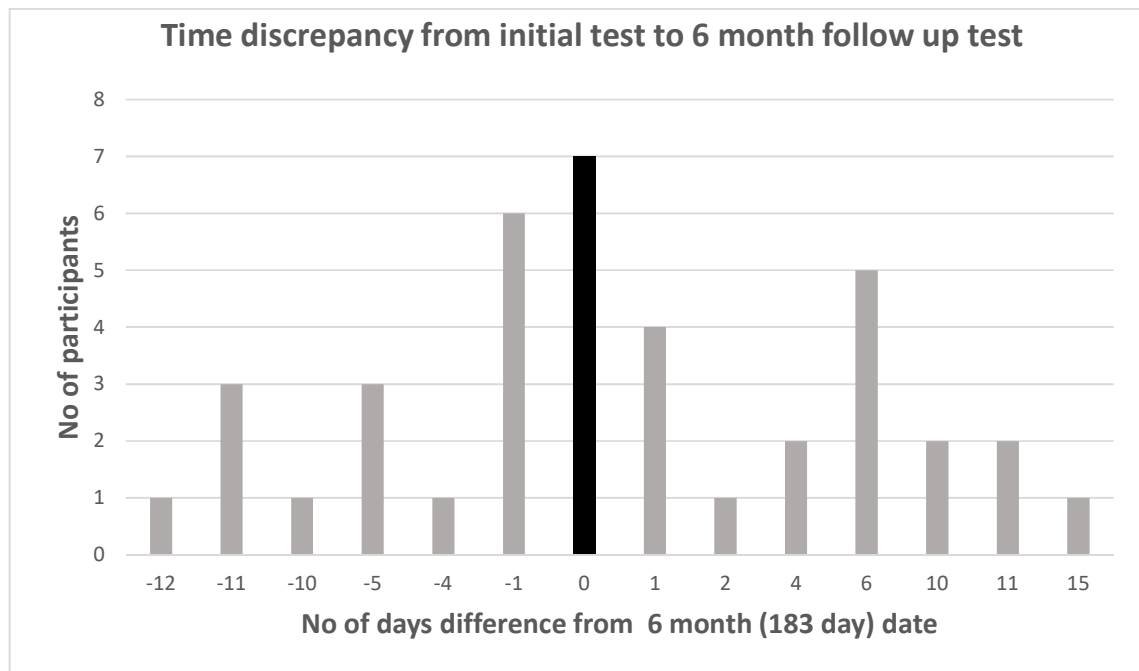
Figure 2. No of days from nominated follow up day at 183 days (marked as zero) and the actual date of test.

The DSDA (iPad version) collates the scores for each test and produces an individual report for each assessment attempt. The DSDA classification, DriveSafe, and DriveAware test scores and the Drive Safe completion time scores were transcribed to the data sheet from the DSDA extended report.

The DSDA classification, recorded on the DSDA extended report, was plotted by the computer program on a nine square matrix. The vertical axis of the matrix was the DriveSafe score (n = 0 to 84). The horizontal axis was the DriveAware score (n = 0 to 17). The DSDA classification was plotted in one of the nine boxes: three green ("*likely to pass an occupational therapy on-road assessment*"), three yellow ("*requires further testing*") or three red boxes ("*likely to fail an occupational therapy on-road assessment*").

For statistical analysis purposes, the boxes were categorised by the researcher (*Figure 3*) as:

- G1, G2, G3 *("likely to pass an occupational therapy on-road assessment"),*
- Y4, Y5, Y6 *("requires further testing"),*
- R7, R8, R9 *("likely to fail an occupational therapy on-road assessment").*
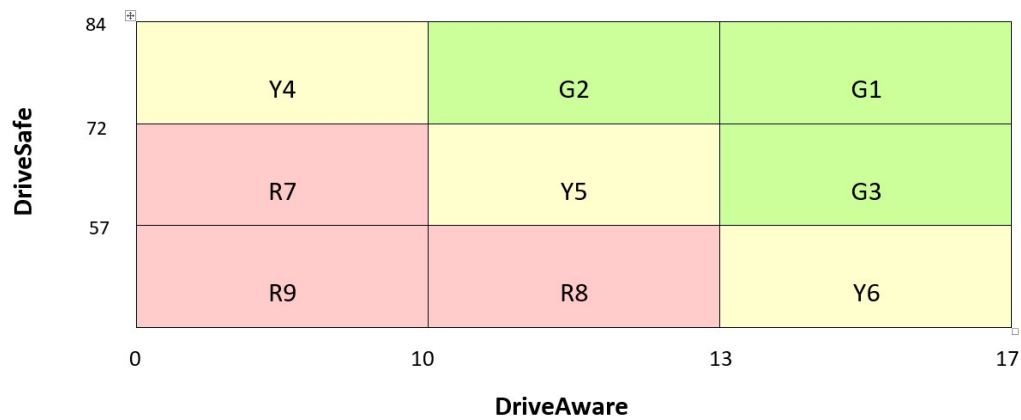
Figure 3. Simulated DSDA classification matrix with rating (G1 to R9) for purposes of statistical analysis. DriveSafe score (0 to 84) on vertical axis and DriveAware score (0 – 17) on horizontal axis.

A high DriveSafe and DriveAware score plots the subject in the top right corner (G1). A low DriveSafe and low Drive Aware score plots the subject in the bottom left corner (R9). The other boxes are variations of these e.g.

- G1 indicates a high DriveSafe (≥72) and DriveAware score (≥ 13),
- G2 Indicates a high DriveSafe Score (≥72) but lower DriveAware score (10 to 12),
- G3 indicates a lower DriveSafe score (57-72) but high DriveAware score (≥ 13).

Despite the scoring differences, G1 to G3 all classify the subject as "*likely to pass an occupational therapy on-road assessment*".

3.4 Statistical analysis

IBM Statistical Package for the Social Sciences (SPSS) Statistics (version 25, IBM Corp.) and OrdinalCont v1.3.0 were used for the statistical analysis of scores reported on the DSDA extended report. Baseline characteristics were grouped into categorical variables. Due to the non-Gaussian distribution of the data, continuous variables were analysed using the non-parametric repeated-measures Friedman test and the post hoc Wilcoxon signed-rank test was used, where required, to distinguish between tests. Mean scores were calculated and compared across the three tests for the entire sample and the discretised groups. The Friedman test assessed differences in scores across the three test attempts with a statistical significance level of 5% and the Wilcoxon signed-rank test assessed differences in scores between the three test attempts with a significance level of 1.7% to allow for a Bonferroni correction. Continuous ordinal regression (97) analysed the relationships between the predictors (age, test number) and the DriveSafe Score. Associations between scores and

variables were assessed using the non-parametric Spearman correlation and parametric Pearson's correlation coefficients. Analysis of variance (ANOVA) was used to assess the influence of age on DS completion time.

The sample was both analysed as a whole and discretised into three groups (young [20 – 49 years]; middle [50 – 69 years]; and elderly [70+ years]). The assumption was made that the young group would have had access to computer technology for most of their lives and would be very familiar with iPad technology. The middle group may not have had computer access in their youth but were likely to have accessed computers and in most cases used iPads. It was hypothesised that participants in the 70+ years age bracket may or may not be familiar with computer technology and were less likely to have regularly used an iPad than the young and middle age participants.

4. Results

4.1 Population Sample

Forty healthy volunteers (17 male and 23 female) were enrolled from the south-eastern region of Australia (*Figure 4*). Each held a current drivers' licence and reported that they regularly drove a car. The median age of the population sample was 58 (range 20 - 91) (*Table 1*).
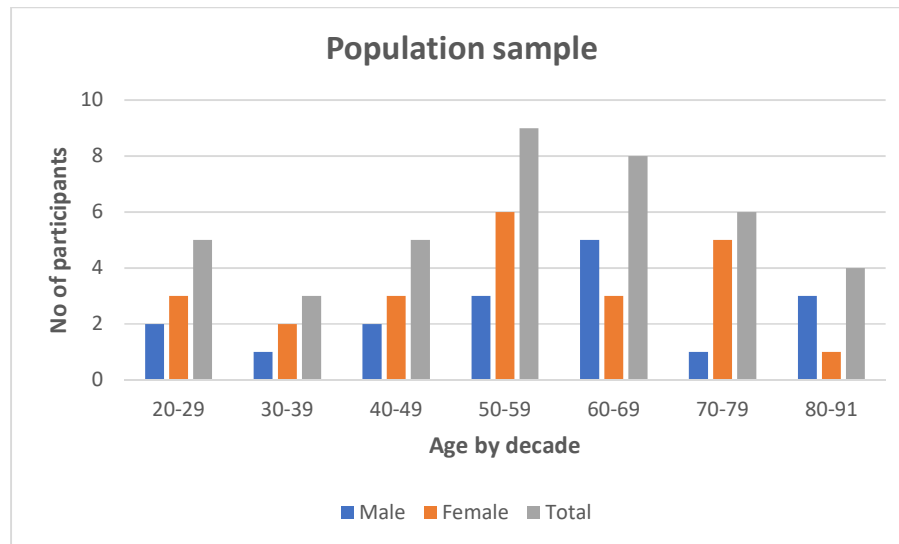
Figure 4. Sample population numbers discretised by age (decades) and gender

Subjects	Median Age (range)	Total
Male	59.0 (20 − 91)	17
Female	57.5 (22 − 81)	23
Total	**58.0 (20 − 91)**	**40**

Table 1. Sample population median age and gender

The sample was discretised into three groups [young (20 − 49 years); middle (50 − 69 years); and elderly (70+ years)] for the purposes of statistical analysis (*Table 2*). The DSDA test, initially administered in July 2017, was repeated in September 2017 (6 weeks after initial testing) and January 2018 (6 months after initial testing).

Subjects	Young	Middle	Elderly	Total
Male	5	8	4*	17*
Female	8	9	6	23
Total	13	17	10*	40*

Table 2. Sample population gender and discretised for age group: young (20 – 49 years); middle (50 – 69 years); and elderly (70+ years)

***One 80-year-old male participant was excluded from the study prior to the third test. His scores from test one and two were included in the analysis.

4.2 DSDA classification

All the participants scored within the G1 to G3 categories (Green – *"likely to pass an on-road occupational therapy assessment"*) for each test; however, some moved between the G1 to G3 categories over the three test attempts.

Between test 1 and 2, five participants changed DSDA classification from G1 to G2; two participants changed from G2 to G1; and one participant changed from G3 to G1.

Between test 2 and 3, five participants changed DSDA classification from G1 to G2; four participants changed from G2 to G1; and one participant changed from G1 to G3.

Between test 1 and 3, five participants changed DSDA classification from G1 to G2 and one participant changed from G2 to G1.

No participants changed from G2 to G3 or G3 to G2.

There was no statistically significant difference in DSDA classification between tests, $\chi^2(2) = 0.5, p = 0.78$.

Change in DSDA classification between tests	Test 1 to Test 2	Test 2 to Test 3	Test 1 to Test 3
G1 to G2	5	5	5
G2 to G1	2	4	1
G1 to G3	0	1	0
G3 to G1	1	0	0
Total no of changes	8	10	6

Table 3. Number of participants who changed classification between tests.

4.3 DriveSafe scores
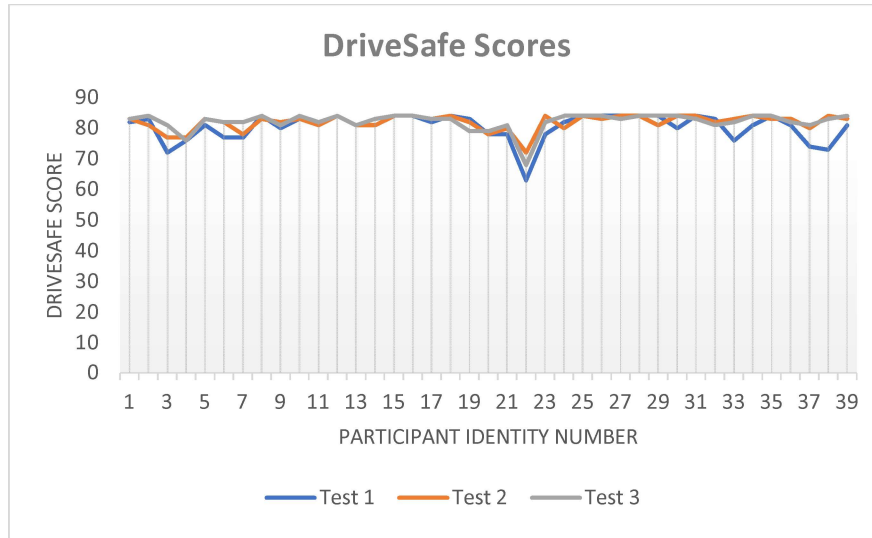
Figure 5. DriveSafe scores comparison for each subject for test 1, 2 and 3

Mean DriveSafe scores (*Table 4*) were high for the study sample (max possible score = 84) but the elderly group had mean scores up to 8 points lower than the young and middle age groups across the 3 tests. The mean scores for the elderly group had a larger standard deviation due to the variation of performance in the elderly age category. All participants were classified as *"likely to pass an on-road assessment"* for every assessment despite differing mean scores on DriveSafe (DS) for different age groups.

Age Group	Mean Test 1 (SD)	Mean Test 2 (SD)	Mean Test 3 (SD)	Mean over 3 tests (SD)
Young (20 – 49)	83.2 (1.16)	83.2 (1.06)	83.6 (0.65)	**83.3 (1.03)**
Middle (50 – 69)	80.9 (3.29)	82.3 (1.72)	82.4 (1.37)	**81.9 (2.35)**
Elderly (70+)	75.1 (5.15)	79.6 (3.65)	79.6 (4.93)	**78.0 (4.95)**
Mean DriveSafe score study sample	**80.2 (4.58)**	**81.9 (2.57)**	**82.2 (2.89)**	**81.4 (3.55)**

Table 4. DriveSafe mean scores and standard deviation for individual tests for each age group test and mean for the three tests (possible score is 0 to 84)

There was a statistically significant difference in DS scores between the three tests, $\chi^2(2) = 9.27$, $p = 0.01$ (Table 5). A Bonferroni correction was applied to all post hoc Wilcoxon signed-rank tests resulting in a significance level set at $p < 0.017$. Post hoc analysis with a Wilcoxon signed–rank test demonstrated a significant difference in DS scores between test 1 and 2

($Z = -2.73, p = 0.006$) and between test 1 and 3 ($Z = -3.36, p = 0.001$). However, there was no significant difference in DS scores between test 2 and 3 ($Z = -0.890, p = 0.37$).

DS Scores	Across 3 tests	Test 1 vs 2	Test 2 vs 3	Test 1 vs 3
Entire Sample	*p = 0.01*	*p = 0.006*	*NSS**	*p = 0.001*
Young (20 – 49)	*NSS*	*NSS*	*NSS*	*NSS*
Middle (50 – 69)	*NSS*	*NSS*	*NSS*	*NSS*
Elderly (70 +)	p = 0.012	p = 0.024**	NSS	p = 0.012

Table 5. Statistical significance of DS scores across all 3 tests and between each test

*NSS = No statistical significance
**Not statistically significant with a Bonferroni correction (p < 0.017) applied but approaching significance

When discretised into age groups there is no statistically significant difference in DS scores over repeated tests in the young (20 – 49 years) age group, $\chi^2(2) = 2.69, p = 0.26.$ Post hoc analysis with a Wilcoxon signed–rank test was conducted and there was no significant difference in DS scores between test 1 and 2 ($Z = -0.276, p = 0.78$), between test 2 and 3 ($Z = -1.5, p = 0.13$), or between test 1 and 3 ($Z = -1.27, p = 0.21$).

The Middle (50 – 69 years) group showed no statistically significant difference between DS scores over repeated tests, $\chi^2(2) = 2.67, p = 0.26.$ Post hoc analysis with a Wilcoxon signed–rank test demonstrated no significant difference in DS scores between test 1 and 2 ($Z = -1.83, p = 0.067$), between test 2 and 3 ($Z = -0.122, p = 0.90$), or between test 1 and 3 ($Z = -1.74, p = 0.081$).

A Friedman test demonstrated a statistically significant difference in DS scores over repeated tests in the elderly (70+ years) group, $\chi^2(2) = 8.88, p = 0.012.$ Post hoc analysis with a Wilcoxon signed–rank test demonstrated a significant difference in DS scores between test 1 and 3 ($Z = -2.52, p = 0.012$). However, there was no significant difference in DS scores between test 1 and 2 ($Z = -2.26, p = 0.024$) or between test 2 and 3 ($Z = -0.303, p = 0.76$). A sensitivity analysis was performed, indicating that a statistically significant difference ($p < 0.017)$ between DS scores in test 1 and 2 would have been reached with the addition of the mean score of just one more subject ($Z = -2.45, p = 0.014$).
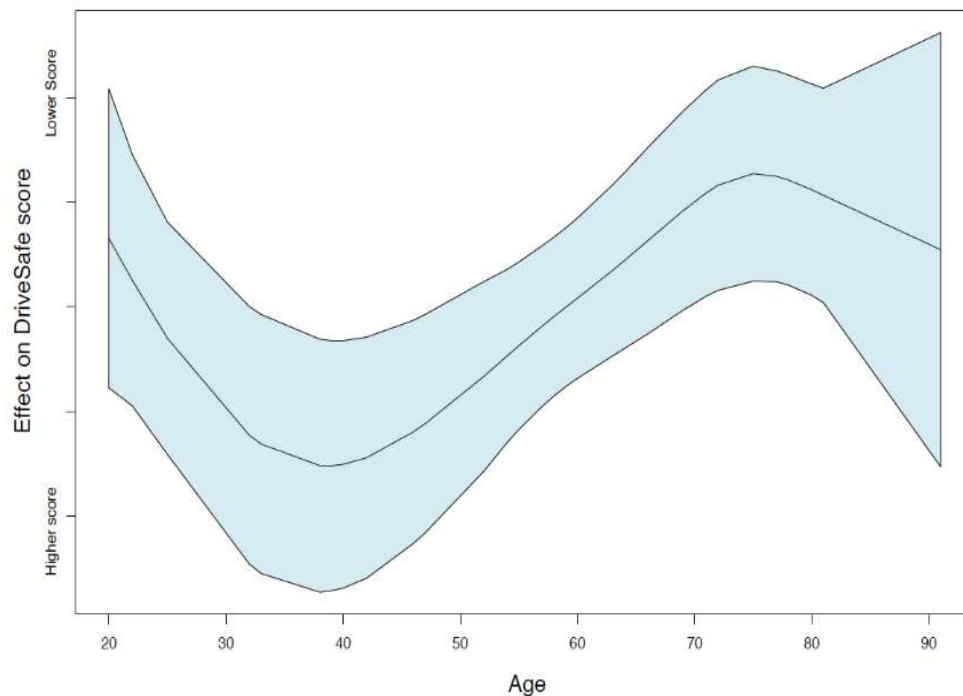
Figure 6. Regression analysis demonstrating relationship DriveSafe scores and age. Top and bottom lines are 95% confidence intervals).

Continuous ordinal regression was used to study the relationship between the predictors (age and test number) and the outcome of interest (DS scores). Age was considered as a continuous scale rather than in the grouped age categories used for the Friedman and Wilcoxon signed-rank test analysis. The effect of repeated measures was taken into consideration by incorporating individual random effects in the model. The analysis demonstrated a relationship between age and DS scores.

Overall mean scores reduced as a healthy subject aged however, the peak performance age in this sample (the age where scores are expected to be highest) was approximately 38 years. Scores from healthy participants over 70 years of age were lower than any other ages 20 to 69 years (Figure 6). The regression analysis approached significance, showing a weak improvement in scores between test 1 and 2 (p-value = 0.08) with an effect size of 0.7 (SD = 0.4) and T value of 1.8, but no difference is demonstrated between test 2 and 3 with this sample size (p-value = 0.31) with an effect size of -0.48 (SD = 0.4) and T value -1.03. It should be noted that in ordinal regression the estimated effects are on the latent scale are not directly on the observed scores. The relation between these two quantities is estimated, and on average in the observed interval (i.e. the range of test scores over the three tests (63-84),

an effect of 1 unit on the latent scale corresponds to a difference of 1.65 on the observed score (the Drive Safe score).

4.3.1 Correlation of DriveSafe scores with gender

A Spearman's correlation assessed the relationship between DS score and gender and determined that there was no statistically significant correlation between DS score and gender. Test 1 $r_s = 0.115, p = 0.48$; test 2 $r_s = -0.007, p = 0.97$; and test 3 $r_s = -0.096, p = 0.56$.

4.4 DriveAware scores

DriveAware (DA) measures the subject's perception of his or her own performance on the DriveSafe test and their current driving performance. It assesses awareness and insight by comparing performance on DriveSafe, to answers to questions both in the DS test and in a clinician interview. Mean scores, on a scale of 0 to 17, are shown in *Table 6.*

Age Group	Mean Test 1 (SD)	Mean Test 2 (SD)	Mean Test 3 (SD)	Mean over 3 tests (SD)
Young (20 – 49)	13.0 (0.8)	13.5 (1.3)	13.3 (1.3)	13.3 (1.1)
Middle (50 – 69)	14.5 (1.3)	13.8 (1.3)	13.9 (1.6)	14.1 (1.4)
Elderly (70+)	15.3 (1.4)	14.6 (1.6)	14.6 (1.6)	14.9 (1.5)
Mean DriveSafe Score study sample	**14.2 (1.5)**	**13.9 (1.4)**	**13.9 (1.5)**	**14 (1.5)**

Table 6. DriveAware mean scores and standard deviation (possible score is 0 to 17)

There was no statistically significant difference in DA score between repeated tests for the study sample, $\chi^2(2) = 3.58, p = 0.17$. Post hoc analysis with Wilcoxon signed–rank tests with a Bonferroni correction applied, resulted in a significance level set at $p < 0.017$. There was no significant difference in DA scores between test 1 and 2 ($Z = -1.37, p = 0.17$), between test 2 and 3 ($Z = -0.047, p = 0.96$), or between test 1 and 3 ($Z = -1.52, p = 0.13$). When discretised, there is no statistically significant difference in DA scores between repeated tests in the young (20 – 49 years) age group, $\chi^2(2) = 0.929, p = 0.63;$ the middle (50 – 69 years) age group, $\chi^2(2) = 4.04, p = 0.13;$ or the elderly (70+ years) age group $\chi^2(2) = 4.22, p = 0.12.$

4.4.1 Correlation of DriveAware scores with gender

A Spearman's correlation assessed the relationship between DA score and gender and determined there was no statistically significant correlation between DA score and gender when the entire population was analysed. Test 1: $r_s = 0.027$, $p = 0.87$; test 2: $r_s = 0.182$, $p = 0.26$; and test 3: $r_s = 0.178$, $p = 0.28$.

The sample was discretised into age groups and a Spearman's correlation was performed to assess the relationship between DA score and gender for test 1, 2 and 3. There was a positive correlation between DA score and gender in the young (20-49 years) group for test 2 and 3, which was statistically significant ($r_s = 0.720$, $p = 0.006$ and $r_s = 0.695$, $p = 0.008$ respectively). Males recorded a lower DA score than females in the young (20 – 49 years) age group on test 2 and 3. There was no statistically significant correlation between DA score and gender for the middle (50 – 69 years) or elderly (70+ years) groups.

4.5 DriveSafe completion time

The computer program records the time the subject takes to complete the DriveSafe section of the test and it is reported on the DSDA extended report. Mean completion times are presented in *Table 7*.

Age Group	Test 1 (SD)	Test 2 (SD)	Test 3 (SD)	Mean across 3 tests (SD)
Young (20 – 49)	7.79 (0.63)	6.40 (0.61)	6.28 (0.72)	6.82 (0.95)
Middle (50 – 69)	8.48 (1.27)	7.35 (1.34)	6.98 (0.84)	7.60 (1.32)
Elderly (70 +)	11.64 (3.40)	10.47 (3.14)	10.18 (4.69)	10.81 (3.71)
Mean Completion time study sample (mins)	**9.09 (2.40)**	**7.80 (2.37)**	**7.49 (2.72)**	**8.13 (2.60)**

Table 7. Mean scores of DriveSafe completion time - the time taken for the population sample and each age group to complete the DriveSafe section of DSDA (minutes)

A Friedman test indicated a statistically significant difference in DS completion times for the whole study population between the three repeated tests, $\chi^2(2) = 0\ 36.4$, $p = 0.001$ with a substantial reduction in completion time over the three tests. Post hoc analysis with Wilcoxon signed–rank test with a Bonferroni correction applied, resulted in a significance level set at $p < 0.017$. There was a statistically significant reduction in DS completion times

between test 1 and 2 ($Z = -4.73, p = 0.001$) and between test 1 and 3 ($Z = -4.87, p = 0.001$). However, there was no significant difference in DS time completion between test 2 and 3 ($Z = -1.90, p = 0.058$) (*Table 8*).

When discretised by age there was a statistically significant difference between DS completion times over repeated tests in the young (20 – 49 years) age group, $\chi^2(2) = 16.6$, $p = 0.001$. Post hoc analysis with a Wilcoxon signed–rank test demonstrated a significant difference in DS completion times between test 1 and 2 ($Z = -3.18, p = 0.001$) and between test 1 and 3 ($Z = -3.11, p = 0.002$), but no significant difference in DS completion times between test 2 and 3 ($Z = -0.594, p = 0.55$).

The middle (50 – 69 years) group demonstrated a statistically significant difference between DS completion times over repeated tests, $\chi^2(2) = 11.8, p = 0.003$. Post hoc analysis with a Wilcoxon signed–rank test demonstrated a significant difference in DS completion times between test 1 and 2 ($Z = -2.39, p = 0.017$) and test 1 and 3 ($Z = -3.39, p = 0.001$), but no significant difference in DS completion times between test 2 and 3 ($Z = -1.52, p = 0.13$).

There was a statistically significant difference in DS completion time for the elderly (70+ years) group over time, $\chi^2(2) = 10.8, p = 0.005$. Post hoc analysis with a Wilcoxon signed–rank test demonstrated a significant difference in DS completion times between test 1 and 2 ($Z = -2.67, p = 0.008$). However, there was no significant difference in DS completion times between test 2 and 3 ($Z = -0.77, p = 0.44$) or between test 1 and 3 ($Z = -1.96, p = 0.051$).

DS Completion time	Across 3 tests	Test 1 vs 2	Test 2 vs 3	Test 1 vs 3
Entire Sample	p = 0.001	p = 0.001	NSS*	p = 0.001
Young (20 – 49)	p = 0.001	p = 0.001	NSS	p = 0.002
Middle (50 – 69)	p = 0.003	P = 0.017	NSS	p = 0.001
Elderly (70+)	p = 0.005	p = 0.008	NSS	NSS

Table 8. Statistical significance of DS completion times across all 3 tests and between each test

4.5.1 Comparison of DriveSafe completion time and age

A Pearson's correlation assessed the relationship between DS completion times and age and demonstrated a strong positive correlation between DriveSafe completion times and age which was statistically significant for test 1 ($r = 0.6$, $n = 40$, $p = 0.001$); test 2 ($r = 0.625$, $n = $

40, $p = 0.001$) and test 3 ($r = 0.512$, $n = 39$, $p = 0.001$). An analysis of variance of DriveSafe completion times showed a statistically significant difference between the completion time means between the young and the elderly age group $(p = 0.001)$ and the middle and elderly age group $(p = 0.001)$; but no statistically significant difference between the young and the middle age group $(p = 0.81)$.

4.5.2 Comparison of DriveSafe completion times and gender

A Pearson's correlation assessed the relationship between DS completion times and gender. There was no correlation between DriveSafe completion times and gender for any of the tests: test 1 ($r = 0.038$, $n = 40$, $p = 0.82$); test 2 ($r = -0.139$, $n = 40$, $p = 0.39$) and test 3 ($r = 0.007$, $n = 39$, $p = 0.97$).

5. Discussion

This study tested the hypotheses that DSDA results of healthy people do not differ over repeated assessments; there is no difference in results between genders over repeated assessments of the DSDA; and there is no difference in results between age groups over repeated assessments with DSDA. The four DSDA components assessed were DSDA classification; DriveSafe scores; DriveAware scores and DriveSafe completion time (elapsed time).

This study aimed to contribute to a larger study identifying a concise functional task assessment that assesses changes in high-level cognition, for repeated administration in a clinic or hospital setting. A literature review identified DriveSafe DriveAware (iPad version) potentially as a suitable assessment tool but despite previously established reliability and validity of DriveSafe DriveAware (DSDA) (87, 98), the test-retest reliability of the iPad version had not yet been established.

The results of this study showed that although there is small learning effect between test 1 and 2 for the DriveSafe score, the scores on all components (DSDA Classification, DriveSafe and DriveAware) remained relatively constant across tests with the exception of the DriveSafe completion time.

Although this study did not aim to assess the validity and reliability of the DSDA touchscreen it should be noted that all the healthy participants in this study who were current drivers were predicted to be "*likely to pass an occupational therapy on-road assessment*" supporting its ability to predict cognitive ability to drive in healthy participants.

5.1 DSDA classification

The DSDA classification is a combination of the DriveSafe score and the DriveAware score plotted on a matrix. Based on these scores, the DSDA assessment classifies the subject into:
1. *"likely to pass an occupational therapy on-road assessment"* (green category 1 to 3);
2. *"requires further testing"* (yellow category 4 to 6); or
3. *"unlikely to pass an occupational therapy on-road assessment"* (red category 7 to 9).

The results of this study supported the hypothesis that DSDA classification score of healthy people does not differ significantly when repeated after 6 weeks and 6 months, demonstrating good test-retest reliability. All the study participants scored within the G1 to G3 categories (Green – *"likely to pass an occupational therapy on-road assessment"*) for all tests (as expected); however, some moved (within the matrix) between the G1 to G3 categories over repeated assessments. Despite movements in the green section of the matrix between G1 to G3, the results were all interpreted in the analysis as *"likely to pass an on-road assessment"*. The G1 to G3 categories were nominated by the researcher for the purpose of this study however the DSDA does not recognise these categories and a participant needs only to fall within one box of the colour classification to achieve the classification.

There was no evidence in this study to assume that a change in DSDA category within the green section is a specific indication of a change in either the subject's cognition or insight/awareness. No participants in this study moved from the green classification to either the yellow or red classification. This is consistent with the predictive strength of the DSDA, as all subjects were healthy and were continuing to safely drive throughout the study.

A change from the green classification section to the yellow or red classification sections of the matrix would indicate less likelihood of passing an on-road driving assessment only; however, in future assessments with brain tumour patients, a clinician may choose to interpret this change as an indication that there has been a possible change in cognition. Results from this study cannot be directly extrapolated to cognitively impaired subjects; however, a change in DSDA classification may be an indicator to a clinician that the patient's cognition or awareness has altered and may prompt further medical or neuropsychological evaluation. Distinguishing whether the patient has changed classification primarily due to poor scores on DriveSafe or due to poor scores on DriveAware may further assist the clinician, as each outcome implies different deficits and causes. However, these suppositions would need to be verified within a future study.

5.2 DriveSafe score

The DriveSafe (DS) score assessed cognition and more specifically memory, attention, and attention to detail. Participants viewed 10 screens of a driving situation and the score

measured how many objects they correctly identified; the correct positions of the objects in a driving scene; and the correct direction of travel. The maximum score was 84 and according to the DSDA matrix, a subject who is "*likely to pass on on-road driving assessment*" should have scored between 58 and 84 depending on their DriveAware score (*Figure 3*). The mean DS score for healthy subjects across repeated tests in this study was 81.4 (SD 3.55). On initial DS testing, the elderly group scored mean scores up to 8 points lower than the middle and young groups respectively; however, by the third attempt the elderly age group mean score was up to 4 points below the middle and young group mean scores respectively. The mean test scores for the elderly group improved by approximately 4 points between test 1 and test 2; however, the mean scores between test 2 and test 3 remained virtually the same (*Table 4*). The elderly group's increased confidence with the previously unfamiliar technology may have led to marginal score improvements.

Continuous ordinal regression demonstrated a relationship between age and DS scores indicating that the peak DS performance in this sample was 38 years. Scores were lower for those over 70 years, revealing reduced DS performance after 70 years of age compared to all other ages. There was a very small improvement between test 1 and 2 scores for the entire sample as the result approached statistical significance (*p = 0.08*) although not between test 2 and 3 suggesting an initial learning effect between the first and second test, albeit a small one.

When the sample was discretised by age this score improvement across the three repeated tests was only statistically significant ($\chi^2(2) = 8.88, p = 0.012$) in the elderly group (70+ years). The difference between test 1 and test 2 for the elderly group appeared initially to approach statistical significance (*p = 0.024*); however, it did not meet the strict threshold required for statistical significance after the application of a Bonferroni correction (resulting in a significance level threshold set at *p < 0.017*). This aligned with the regression analysis. A sensitivity analysis indicated that statistical significance would have been reached with the addition of just one more individual. This implied a small change between test 1 and 2 for the elderly group supporting good test-retest reliability. There was no significant difference in scores between test 2 and 3 for the elderly group.

Learning effects have been demonstrated previously in test-retest studies of cognitive tests (89, 90) most notably, between the first two tests of a series. Test-retest studies of vision in normal individuals and patients with eye disease (99-102), tests of motor function in healthy individuals and in patients with multiple sclerosis and Down syndrome (103-105), tests of cognition in normal individuals and patients following subarachnoid haemorrhage (106), and tests of concussion in children and adolescents (107-109) also report learning effects. In many of the published studies of repeated testing, the learning effect occurred between the first and second episode of testing but disappeared on subsequent testing (90, 99-102, 104, 105).

One possible explanation for the findings in this study was that the small learning effect associated with the improvement in scores demonstrated in the elderly group between test 1 and 2 was related to initial unfamiliarity with the iPad technology and anxiety regarding the possibility of being assessed as unfit to drive. On subsequent attempts, the participants were more familiar with the technology and displayed less anxiety, presumably because they knew that they had "passed" the test the first time and knew they were unlikely to "fail" on later attempts. Like previous test-retest studies (90, 99-102, 104, 105), after the improvement between the first and second test, the 70+ years participants showed no significant change to their scores analogous to the participants aged 20 to 69 years. Future studies may expect a small DS score improvement in healthy subjects between the first and second assessment and this improvement is likely to be more evident in people over 70 years of age.

Although the results of this study cannot be directly related to populations with brain tumour or cognitive impairment, significant increases or decreases (beyond the standard deviation of 3.55) in the DriveSafe score between tests (especially after test 2 for those over 70 years) may be an indicator to a clinician that the patient's cognition has altered and may prompt further medical or neuropsychological evaluation.

5.2.1 Ceiling effect

It was difficult to determine whether there would have been a greater improvement in the young and middle age healthy participants because of the possible ceiling effect of the scoring. Twelve of the 30 participants from 20 to 69 years of age achieved the maximum

score of 84 for test 1. Only two-thirds of these participants (n = 8) repeated this result on test 2. For the one-third (n = 4) who did not score 84 on test 2, scores reduced minimally (by one to two points). Only five participants (12.5% of the entire sample, n = 40) recorded a top score of 84 for all three tests, suggesting that the ceiling effect may not be substantial. However, a larger sample would be required to confirm this.

5.2.2 DriveSafe and gender

There was no association between DS score and gender for any of the three age groups as measured by a Spearman correlation.

5.3 DriveAware score

DriveAware (DA) assessed insight and awareness, an essential characteristic for safe driving. Based on the subject's perception, it compared the DS score to responses from the subject about their performance on the DS test and their driving ability. A subject may have scored lower on the DS, but a high DA score denoted that they may still be classified as "*likely to pass an occupational therapy on-road assessment*". Drivers have been reported to take compensatory driving and safety actions when aware of reduced cognitive functioning (98) e.g. no longer driving at night or avoiding peak hour driving. The maximum score is 17 and according to the DSDA matrix, a subject who is "*likely to pass on on-road driving assessment*" should have scored between 11 and 17 depending on their DS score (*Figure 3*). The mean score across all three tests in this study of healthy subjects was 14 (SD 1.5).

This study indicated that the DA score remained relatively consistent for healthy subjects (within one or two points - which is close to the standard deviation of 1.5) across three repeated test attempts supporting DA test-retest reliability in healthy subjects. DA scores generally remained consistent but a small reduction (1 to 2 points less) on subsequent tests may have indicated more confidence in ability to complete the DS test successfully. Given the confirmed small improvement in DS scores between test 1 and 2, this slight reduction in DA score and increased confidence, was to be expected.

The mean DA scores of the young age group were approximately one point less, and the elderly group, one point more, than the sample mean score. The DA mean scores were slightly lower for the young age group but not statistically significant. The young age group

subjects may have perceived that their performance (DS score) was good, resulting in lower DA scores. Conversely, the elderly group may have perceived their DS scores as "OK/Not too bad" or "not so good", resulting in higher DA scores. The higher DA scores in the elderly group suggested they were aware that they might have missed some of the DS answers. This may have been a confirmation that healthy elderly drivers were more aware of limitations with their driving; e.g. the researcher noted in response to the question, "How often do vehicles or pedestrians 'appear out of nowhere'?" the elderly group were more likely to answer, "occasionally". This answer, admitting fallibility and awareness of such, consequently raised their DriveAware score. Although unable directly to infer that self-regulation exists, Coxon et al. (44) raised self-regulation in the older driver "as a promising strategy for preserving safe mobility and independence in later life".

Significant increases or decreases in DA score (beyond 4 points – which is more than double the standard deviation) may be a potential indicator to a clinician that the patient's cognition has altered and may trigger further medical or neuropsychological evaluation. Moreover, a consistent DA score combined with a significant reduction in the DS score may indicate decreased awareness or insight in reduced performance and prompt further investigation. However, this conjecture would need to be verified in a future study.

5.3.1 DriveAware and gender

There was no significant correlation between DA Scores and gender on analysis of the whole sample of healthy subjects. There was a correlation between males in the young (20 – 49 years) group and a slightly lower DA score combined with consistently high DS scores. Interestingly, the young age group women also had high DS scores but higher DA scores than the men, suggesting less confidence in their performance on the test. This may be consistent with findings that young men demonstrate more confidence in their capabilities than young women despite a similar level of performance (110).

5.4 DriveSafe completion time

The DS completion time appeared on the DS extended report and specified the time the subject took to complete the DS test. Most of the study sample completed the DS test in 10 minutes or less. With the addition of the DriveAware clinician interview, DSDA completion time was approximately 15 to 25 minutes. This contrasted with more time-consuming

neuropsychological testing where a two-and-a-half-hour assessment may be considered a short assessment (111).

The DS completion time revealed significant variability on repeated tests and did not demonstrate test-retest reliability for healthy subjects. Most participants completed the second test in a significantly shorter time than the first attempt suggesting a substantial learning effect. Overall, younger participants completed the test in significantly less time, with elderly participants taking more time. There was an improvement in DS completion time between the three tests, for the whole sample ($\chi2(2) = 36.4, p = 0.001$). This effect was significant between test 1 and 2, but not between test 2 and 3.

The DSDA extended report states that "research indicates that the median time taken to complete DriveSafe for people who pass an occupational on-road assessment is 4 minutes and 46 seconds, and for people who fail the assessment, 6 minutes and 49 seconds". None of the participants in this study completed the DriveSafe in 4 minutes and 46 seconds. The fastest completion time was 5 minutes and 9 seconds achieved by a 58-year-old female participant on her second attempt however her time increased by 20 seconds on her third attempt. The mean time for all participants for all tests was 8.13 minutes (8 minutes 8 seconds) (SD 2.6) so the results of this study of healthy participants did not corroborate the time measure advice on the extended report.

The researcher observed that the DS completion time recorded on the DSDA extended report appeared to include the time taken to watch the instructions and complete the three practice screens in addition to the 10 screens of the test. This may explain the discrepancy between the completion times in this study and the DSDA time range on the DSDA extended report product information that indicates the expected time range for a subject if they are "*likely to pass an on-road assessment".* It was noted that the younger age group, and to a lesser extent the middle age group, rushed through the instructions and practice screens on repeated attempts. The elderly group perhaps less familiar with the iPad technology and less confident in their performance, still fastidiously completed the instructions and all practice screens on subsequent tests. The completion time reduction that is more marked in the younger and middle age groups but less evident in the elder group may be explained by this rationale.

The DS completion time was not a reliable measure over repeated tests in its current form, in that it was not consistent between assessments and did not meet the completion time recommendation on the DSDA extended report. However, it may still be a useful measure of performance. This study indicated that healthy subjects could be expected to significantly reduce their DS completion time on repeated testing (as measured by the time on the DSDA extended report), especially those in the younger and middle age groups. Further medical or neuropsychological evaluation may be warranted if patient performance declines (i.e. they take longer to complete the DS test).

5.5 Strengths of the study

The strengths of the study included participant inclusion and exclusion criteria, the strict assessment schedule and the non-disclosure of results to the participants. The sample represented a large section of the Australian population and included men and women aged 20 to 91 years of age. Highest level of education completed ranged from year 10 high school to Master's degree. Participants were born in seven different countries, but all spoke English proficiently. All participants held, at minimum, a class C drivers' licence and reported that they drove regularly without major concerns.

The inclusion and exclusion criteria controlled the possibility of cognitive changes affecting results. One male 80-year-old participant was excluded from the study prior to test three under the criteria; however, his results from the first two assessments were retained in the data and included in the assessment of score changes between test 1 and 2 but not between test 1 and 3 or test 2 and 3. Although it may have been interesting to complete test 3 on this subject (and determine whether he was affected by medication usage), it was not within the aim of this study to investigate the effect of medication on DSDA.

The follow-up six-week test occurred within one week of the six-week date and the six-month test was performed within two weeks of the six-month date from the initial test. This aligned with the variability expected with patient clinic appointments.

All tests were administered by the researcher (a registered occupational therapist) using the same equipment to eliminate further variables, and although participants were informed whether they had met the criterion of "*likely to pass an on-road assessment*", all participants

were blinded to their DS and DA test scores and their DS completion time scores. Participants received no information during the study as to whether their scores had improved or deteriorated.

5.6 Study Limitations

The limitations of the study included a lack of patient diversity – participants were cognitively fit, English speaking subjects only; lower participant numbers in the 70+ age group; and repeating the test over 6 months only rather than 12 months.

A key limitation of the study was that the sample was limited to healthy, cognitively intact participants despite the DSDA being designed to assess the ability to drive of people with cognitive impairment. Therefore, results cannot be directly translated to diseased or cognitively impaired populations. All participants were classified in the green "*likely to pass an occupational therapy on-road assessment*" DSDA classification, and therefore this study did not assess test-retest reliability for those who score in the yellow or red classifications. Resources were not available during this research project to independently assess and monitor cognitive changes (either improvement or deterioration) in subjects with cognitive deficits, or to undertake a study directly comparing DSDA against the gold standard for cognitive assessment, such as a battery of detailed neuropsychological testing. Testing the same cognitively impaired sample repeatedly with the DSDA assessment and an on-road occupational therapy driving assessment would also have enabled a direct comparison; however, such a comparison was not the primary aim of this study.

The sample was a convenience sample and sample size was determined by the ICC from the test-retest study on a previous version of DSDA.

The sample was limited to participants who spoke English proficiently as the DSDA iPad version was not validated for use with non-English speakers (1). As such, the results are not generalisable to non-English speaking populations. The researcher also had difficulty recruiting healthy volunteers in the 70+ age range who still held a valid drivers' licence and regularly drove an automobile. Future research should aim to recruit more participants in this age group, so the elderly group has similar numbers to the middle age group. Lower

participant numbers in the elderly group meant that one outlier result could affect statistical results significantly. Reporting of median scores rather than mean scores may counter this.

Due to the project time constraints, repeated testing occurred at six weeks and six months only. Repeating the assessment at 12 months would enable comparison of data with patients assessed repeatedly for 12 months in the clinic.

Another key limitation to the study relates to the DS completion time. It appeared to the researcher that the time reported on the DSDA extended report included the time taken to complete the practice screens and none of the healthy participants completed the test near the time that the test publisher reported as the median time achieved by participants who "passed" the test. A future study should confirm this assertion. The researcher does not recommend the DS Completion time, in its current form, as a reliable measure of performance improvement. If utilised, however, the expectation would be that healthy subjects would reduce their time to complete the assessment substantially after the first assessment.

6. Conclusions

This study aimed to demonstrate the test-retest reliability of the DSDA iPad version. This was as a part of a larger study to identify a concise functional outcome measure of high-level cognition suitable to use in repeated testing with brain tumour patients in a clinic or hospital setting. Deterioration in cognition may be an early sign of tumour recurrence after treatment and cognitive testing may provide this alert before radiological evidence (19). It may also be a valuable predictor of survival in some forms of brain tumour (20, 21). It was essential that the identified measure demonstrated test-retest reliability to realise the intention of using it at repeated clinic visits to assess changes in cognition in addition to predicting ability to drive. DriveSafe DriveAware (DSDA) is an assessment developed for the purpose of distinguishing whether people are: *"likely to pass an occupational therapy on-road assessment"*; *"unlikely to pass an occupational therapy on-road assessment"*; or whether the result is unclear and *"further testing"* is required. It assesses cognitive ability to drive, and insight and awareness into ability to drive. Driving is a key activity for independent function, and as such, is highly valued by patients (45).

This study has demonstrated that when used on healthy subjects, the DSDA (iPad version) is a reliable method for repeated assessment, albeit with a small learning effect affecting the DS score between the first and second testing (but not subsequent testing). This effect manifests more noticeably in subjects over 70 years of age and is most likely because of increasing familiarity with the technology and decreasing anxiety over repeated assessments rather than because of improved cognition. The results also highlighted that DS scores are lower overall for people over 70 years of age but still within the cut off limits reported for the DSDA (1). This suggests to clinicians seeking to use this assessment in the future to monitor cognition that lower DS scores are to be expected in those aged 70+ years and are not an indication of any cognitive changes other than those related to age. In addition, DSDA classification and DriveAware scores of healthy subjects showed no statistically significant difference over time, supporting a conclusion of good test-retest reliability for these measures.

In contrast, the DS completion time is not a consistent measure, as all groups reduced their time to complete the test over subsequent tests.

DSDA was developed to assess cognitive ability to drive however it may be useful for identifying cognitive changes in brain tumour patients not detected by other self-care assessments. Although results from this study cannot be directly extrapolated to cognitively impaired subjects, a substantial change in DSDA classification, DS Score or DA score may be an indicator to a clinician that a patient's cognition or awareness has altered and may prompt further medical or neuropsychological evaluation.

7. Future Research Recommendations

Considering the conclusions of this study the following recommendations for future research are made:

1. Repeat the DSDA iPad version assessment with healthy subjects from this study 12 months after their initial assessment.

2. Assess the suitability of using DriveSafe DriveAware (DSDA) to assess high-level cognition of patients with brain tumours.

3. Reassess the DriveSafe completion time measure to determine whether the practice screens are included in the time calculation.

4. Compare the DSDA iPad version results of subjects diagnosed with an illness or condition that affects cognition to a reliable and valid neuropsychological assessment with demonstrated test-retest reliability.

References

1.   Cheal B, Kuang H. DriveSafe DriveAware for Touch Screen: A Screening Tool for Cognitive Fitness to Drive - Administration Manual. Australia: Pearson Education, Pearson Australia Group Pty Ltd; 2015. 52 p.

2.   Department of Health 2018. Brain Cancer. Canberra: Australian Government. Viewed 1 May 2018, <available from: http://www.health.gov.au/internet/main/publishing.nsf/content/A091E85A15E4F93DCA258 14E0005C8EE/$File/brain-cancer-fact-sheet-1.pdf>.

3.   Liu R, Page M, Solheim K, Fox S, Chang SM. Quality of life in adults with brain tumors: Current knowledge and future directions. Neuro-Oncology2009. p. 330-9.

4.   Taphoorn M, Sizoo E, Bottomley A. Review on quality of life issues in patients with primary brain tumors. Oncologist. 2010;15(6):618-26.

5.   Cancer Institute NSW: NSWOG for Neuro-oncology, n.d., Brain Tumours and Driving A guide for clinicians, Cancer Council NSW.

6.   Australian Institute of Health and Welfare 2017. Brain and other central nervous system cancers. Cat. no. CAN 106. Canberra: AIHW.

7.   Ostrom QT, Gittleman H, Liao P, Vecchione-Koval T, Wolinsky Y, Kruchko C, et al. CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010-2014. Neuro-oncology. 2017;19(suppl_5):v1.

8.   Veretennikoff K, Walker D, Biggs V, Robinson G. Changes in Cognition and Decision Making Capacity Following Brain Tumour Resection: Illustrated with Two Cases. Brain Sciences. 2017;7(10):122.

9.   Henriksson R, Asklund T, Poulsen H. Impact of therapy on quality of life, neurocognitive function and their correlates in glioblastoma multiforme: a review. Journal of Neuro-Oncology. 2011;104(3):639-46.

10.   Smith-Cohn M, Chen Z, Peereboom D, Stevens G. Maximizing Function and Quality of Life of Patients with Glioblastoma after Surgical Resection: A Review of Current Literature. Journal of Cancer Therapy. 2016;7(12):857.

11.   Hunt, G (Minister for Health) 2017, $100 million fund to fight brain cancer, media release, Department of Health, Commonwealth of Australia, 29 October, viewed 19 November 2017, <Available from: <http://www.health.gov.au/internet/ministers/publishing.nsf/Content/health-mediarel-yr2017-hunt111.htm>

12.   Weller M, van Den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. The Lancet Oncology. 2017;18(6):e315-e29.

13.   Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. The New England Journal of Medicine. 2005;352(10):987-96.

14.   Soffietti R, Abacioglu U, Baumert B, Combs S, Kinhult S, Kros J, et al. Diagnosis and treatment of brain metastases from solid tumors: guidelines from the European Association of Neuro-Oncology (EANO). Neuro-Oncology. 2017;19(2):162-74.

15.   Sampson JH, Maus MV, June CH. Immunotherapy for Brain Tumors. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(21):JCO2017728089.

16.   Schagen SB, Klein M, Reijneveld JC, Brain E, Deprez S, Joly F, et al. Monitoring and optimising cognitive function in cancer patients: Present knowledge and future directions. European Journal of Cancer Supplements. 2014;12(1):29-40.

17.   Taphoorn M, Klein M. Cognitive deficits in adult patients with brain tumours. The Lancet Neurology. 2004;3(3):159-68.

18.   Gehrke AK, Baisley MC, Sonck ALB, Wronski SL, Feuerstein M. Neurocognitive deficits following primary brain tumor treatment: systematic review of a decade of comparative studies. Journal of Neuro-Oncology. 2013;115(2):135-42.

19.   Meyers CA, Hess KR. Multifaceted end points in brain tumor clinical trials: cognitive deterioration precedes MRI progression. Neuro-oncology. 2003;5(2):89.

20.   Klein M, Postma TJ, Taphoorn MJB, Aaronson NK, Vandertop WP, Muller M, et al. The prognostic value of cognitive functioning in the survival of patients with high-grade glioma. Neurology. 2003;61(12):1796-8.

21.   Johnson DR, Sawyer AM, Meyers CA, O'Neill BP, Wefel JS. Early measures of cognitive function predict survival in patients with newly diagnosed glioblastoma. Neuro-Oncology. 2012;14(6):808-16.

22.   Driscoll JJ, Rixe O. Overall Survival: Still the Gold Standard: Why Overall Survival Remains the Definitive End Point in Cancer Clinical Trials. The Cancer Journal. 2009;15(5):401-5.

23.   Jackson JM, Scheid K, Rolnick SJ. Development of the cancer survivorship care plan: what's next? Life after cancer treatment. Clinical journal of oncology nursing. 2013;17(3):280.

24.   Dirven L, Reijneveld J, Aaronson N, Bottomley A, Uitdehaag B, Taphoorn M. Health-related Quality of Life in Patients with Brain Tumors: Limitations and Additional Outcome Measures. Current Neurology and Neuroscience Reports. 2013;13(7):1-9.

25.   Wen PY, Chang SM, Van Den Bent MJ, Vogelbaum MA, Macdonald DR, Lee EQ. Response Assessment in Neuro-Oncology Clinical Trials. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(21):JCO2017727511.

26.   Ellingson BM, Wen PY, Cloughesy TF. Modified Criteria for Radiographic Response Assessment in Glioblastoma Clinical Trials. Neurotherapeutics. 2017.

27.   FDA. Clinical Outcome Assessment (COA): Glossary of Terms Maryland, USA: U.S. Food & Drug Administration; 2017 [updated 08/23/2017]. Viewed 16 November 2017. Available from: https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm.

28.   Salvo NJ, Hadi S, Napolskikh J, Goh P, Sinclair E, Chow E. Quality of life measurement in cancer patients receiving palliative radiotherapy for symptomatic lung cancer: A literature review. Current Oncology. 2009;16(2):13.

29.   McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs. Medical Care. 1993;31(3):247-63.

30.   Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. JNCI: Journal of the National Cancer Institute. 1993;85(5):365-76.

31.   Webster K, Cella D, Yost K. The F unctional A ssessment of C hronic I llness T herapy (FACIT) Measurement System: properties, applications, and interpretation. Health and Quality of Life Outcomes. 2003;1(1):79.

32.   Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky performance status scale: An examination of its reliability and validity in a research setting. Cancer. 1984;53(9):2002-7.

33.   Wilson H, Dashiell-Aje E, Anatchkova M, Coyne K, Hareendran A, Leidy NK, et al. Beyond study participants: a framework for engaging patients in the selection or development of clinical outcome assessments for evaluating the benefits of treatment in medical product development. Quality of Life Research. 2017.

34.   Wefel JS, Vardy J, Ahles T, Schagen SB. International Cognition and Cancer Task Force recommendations to harmonise studies of cognitive function in patients with cancer. The Lancet Oncology. 2011;12(7):703-8.

35.   Olson R, Tyldesley S, Carolan H, Parkinson M, Chhanabhai T, McKenzie M. Prospective comparison of the prognostic utility of the Mini Mental State Examination and the Montreal Cognitive Assessment in patients with brain metastases. Supportive Care in Cancer. 2011;19(11):1849-55.

36.   Naehrig DN, Koh ES, Vogiatzis M, Yanagisawa W, Kwong C, Shepherd HL, et al. Impact of cognitive function on communication in patients with primary or secondary brain tumours. J Neurooncol. 2016;126(2):299-307.

37.   Robinson GA, Biggs V, Walker DG. Cognitive Screening in Brain Tumors: Short but Sensitive Enough? Frontiers in Oncology. 2015;5(60).

38.   Brazil L, Thomas R, Laing R, Hines F, Guerrero D, Ashley S, et al. Verbally administered Barthel Index as functional assessment in brain tumour patients. Journal of Neuro-Oncology. 1997;34(2):187-92.

39.   Greenberg E, Treger I, Ring H. Rehabilitation outcomes in patients with brain tumors and acute stroke: comparative study of inpatient rehabilitation. Am J Phys Med Rehabil. 2006;85(7):568-73.

40.   Nott MT, Chapparo C, Heard R. Reliability of the Perceive, Recall, Plan and Perform System of Task Analysis: A criterion-referenced assessment. Australian Occupational Therapy Journal. 2009;56(5):307-14.

41.   Lewis J, Chapparo C, Mackenzie L, Ranka J. Work after breast cancer: Identification of cognitive difficulties using the Perceive, Recall, Plan, and Perform (PRPP) System of Task Analysis. British Journal of Occupational Therapy. 2016;79(5):323-32.

42.   WHO. ICF : International classification of functioning, disability and health / World Health Organization. Geneva: World Health Organization; 2001..

43.   Dickerson AE. Driving Assessment Tools Used by Driver Rehabilitation Specialists: Survey of Use and Implications for Practice. American Journal of Occupational Therapy. 2013;67(5):564-73.

44.   Coxon K, Chevalier A, Lo S, Ivers R, Brown J, Keay L. Behind the Wheel: Predictors of Driving Exposure in Older Drivers. Journal of the American Geriatrics Society. 2015;63(6):1137-45.

45.   Dickerson AE, Reistetter T, Gaudy JR. The Perception of Meaningfulness and Performance of Instrumental Activities of Daily Living From the Perspectives of the Medically At-Risk Older Adults and Their Caregivers. Journal of Applied Gerontology. 2013;32(6):749-64.

46.   Field KM, Guyatt N, Fleet M, Rosenthal MA, Drummond K. Communicating with brain tumor patients about driving: Are we falling short of the mark? Asia-Pacific Journal of Clinical Oncology. 2014;10(1):92-3.

47.   Mansur A, Hird M, Desimone A, Pshonyak I, Schweizer T, Das S. Driving habits and behaviors of patients with brain tumors: a self-report, cognitive and driving simulation study. Sci Rep. 2018;8(1):4635-.

48.   Betz ME, Jones J, Carr DB. System facilitators and barriers to discussing older driver safety in primary care settings. Injury Prevention. 2015;21(4):231.

49.   CCA (Cancer Council Australia) 2016. Understanding brain tumours: a guide for people with brain or spinal cord tumours, their families and friends. Sydney: CCA. Viewed 27 October 2017. Available from: <https://www.cancercouncil.com.au/wp-content/uploads/2014/05/CAN879-Brain-NSW.pdf>.

50.	Chin YS, Jayamohan J, Clouston P, Gebski V, Cakir B. Driving and patients with brain tumours: a postal survey of neurosurgeons, neurologists and radiation oncologists. Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia. 2004;11(5):471-4.

51.	Gaertner J, Vent J, Greinwald R, Rothschild MA, Ostgathe C, Kessel R, et al. Denying a patient's final will: public safety vs. medical confidentiality and patient autonomy. J Pain Symptom Manage. 2011;42(6):961-6.

52.	Thomas S, Mehta MP, Kuo JS, Ian Robins H, Khuntia D. Current practices of driving restriction implementation for patients with brain tumors. J Neurooncol. 2011;103(3):641-7.

53.	McDonald KL. Brain tumors and driving. J Neurooncol. 2011;104(1):399-400.

54.	Austroads, Commission NT. Assessing Fitness to Drive for commercial and private vehicle drivers A resource for health professionals in Australia. 5th ed. Sydney: Austroads Ltd; 2016. 172 p.

55.	CCA, Australian Cancer Network & COSA 2009. Clinical practice guidelines for the management of adult gliomas: astrocytomas and oligodendrogliomas. 1st edn. Sydney: CCA.

56.	Dickerson AE, Brown Meuel D, Ridenour CD, Cooper K. Assessment Tools Predicting Fitness to Drive in Older Adults: A Systematic Review. American Journal of Occupational Therapy. 2014;68(6):670-9.

57.	Kay LG, Bundy AC, Clemson LM. Predicting Fitness to Drive in People With Cognitive Impairments by Using DriveSafe and DriveAware. Archives of Physical Medicine and Rehabilitation. 2009;90(9):1514-22.

58.	Kay LG, Bundy AC, Clemson L, Cheal B, Glendenning T. Contribution of off-road tests to predicting on-road performance: A critical review of tests. Australian Occupational Therapy Journal. 2012;59:89-97.

59.	Gibbons C, Smith N, Middleton R, Clack J, Weaver B, Dubois S, et al. Using Serial Trichotomization With Common Cognitive Tests to Screen for Fitness to Drive. American Journal of Occupational Therapy. 2017;71(2):1-8.

60.	Krasniuk S, Classen S, Morrow SA, Monahan M, Danter T, Rosehart H, et al. Driving Errors That Predict On-Road Outcomes in Adults With Multiple Sclerosis. OTJR: Occupation, Participation and Health. 2017;37(3):115-24.

61.	Allan C, Coxon K, Bundy A, Peattie L, Keay L. DriveSafe and DriveAware Assessment Tools Are a Measure of Driving-Related Function and Predicts Self-Reported Restriction for Older Drivers. Journal of Applied Gerontology. 2016;35(6):583-600.

62.	Bédard M, Campbell S, Riendeau J, Maxwell H, Weaver B, Bédard M. Visual-cognitive tools used to determine fitness-to-drive may reflect normal aging. Clinical & Experimental Optometry. 2016;99(5):456-61.

63.	Bennett JM, Chekaluk E, Batchelor J. Cognitive Tests and Determining Fitness to Drive in Dementia: A Systematic Review. Journal of the American Geriatrics Society. 2016;64(9):1904-17.

64.	Piersma D, Fuermaier ABM, de Waard D, Davidse RJ, de Groot J, Doumen MJA, et al. Prediction of Fitness to Drive in Patients with Alzheimer's Dementia. PLOS ONE. 2016;11(2):e0149566.

65.	Baker A, Unsworth CA, Lannin NA. Determining fitness to drive: A systematic review of the methods and assessments used after mild traumatic brain injury. British Journal of Occupational Therapy. 2015;78(2):73-84.

66.	Hollis AM, Duncanson H, Kapust LR, Xi PM, O'Connor MG. Validity of the Mini-Mental State Examination and the Montreal Cognitive Assessment in the Prediction of Driving Test Outcome. Journal of the American Geriatrics Society. 2015;63(5):988-92.

67.	Barco PP, Wallendorf MJ, Snellgrove CA, Ott BR, Carr DB. Predicting Road Test Performance in Drivers With Stroke. American Journal of Occupational Therapy. 2014;68(2):221-9.

68.	Cullen N, Krakowski A, Taggart C. Early neuropsychological tests as correlates of return to driving after traumatic brain injury. Brain Injury. 2014;28(1):38-43.

69.	Hines A, Bundy AC. Predicting driving ability using DriveSafe and DriveAware in people with cognitive impairments: A replication study. Australian Occupational Therapy Journal. 2014;61(4):224-9.

70.	Motta K, Lee H, Falkmer T. Post-stroke driving: Examining the effect of executive dysfunction. Journal of Safety Research. 2014;49(Supplement C):33.e1-8.

71.	Bowers AR, Anastasio RJ, Sheldon SS, O'Connor MG, Hollis AM, Howe PD, et al. Can we improve clinical prediction of at-risk older drivers? Accident Analysis & Prevention. 2013;59(Supplement C):537-47.

72.	Classen S, Wang Y, Winter SM, Velozo CA, Lanford DN, Bedard M. Concurrent criterion validity of the safe driving behavior measure: a predictor of on-road driving outcomes. AJOT: American Journal of Occupational Therapy. 2013;67(1):108.

73.	Dobbs BM, Shergill SS. How effective is the Trail Making Test (Parts A and B) in identifying cognitively impaired drivers? Age & Ageing. 2013;42(5):577-81.

74.	Hoggarth PA, Innes CRH, Dalrymple-Alford JC, Jones RD. Predicting On-Road Assessment Pass and Fail Outcomes in Older Drivers with Cognitive Impairment Using a Battery of Computerized Sensory-Motor and Cognitive Tests. Journal of the American Geriatrics Society. 2013;61(12):2192-8.

75.	Marino M, de Belvis A, Basso D, Avolio M, Pelone F, Tanzariello M, et al. Interventions to evaluate fitness to drive among people with chronic conditions: Systematic review of literature. Accident Analysis & Prevention. 2013;50(Supplement C):377-96.

76. Ott BR, Davis JD, Papandonatos GD, Hewitt S, Festa EK, Heindel WC, et al. Assessment of Driving-Related Skills Prediction of Unsafe Driving in Older Adults in the Office Setting. Journal of the American Geriatrics Society. 2013;61(7):1164-9.

77. Hargrave DD, Nupp JM, Erickson RJ. Two brief measures of executive function in the prediction of driving ability after acquired brain injury. Neuropsychological Rehabilitation. 2012;22(4):489-500.

78. Kay LG, Bundy AC, Cheal B. DriveSafe and DriveAware: A promising new off-road test to predict on-road performance. In: Dorn L, editor. Driver Behaviour and Training. Ashgate: Cranfield University; 2012. p. 127-40.

79. Unsworth CA, Baker A, Taitz C, Chan SP, Pallant JF, Russell KJ, et al. Development of a standardised Occupational Therapy - Driver Off-Road Assessment Battery to assess older and/or functionally impaired drivers. Australian Occupational Therapy Journal. 2012;59(1):23-36.

80. Bliokas VV, Taylor JE, Leung J, Deane FP. Neuropsychological assessment of fitness to drive following acquired cognitive impairment. Brain Injury. 2011;25(5):471-87.

81. Fields SM, Unsworth CA. Revision of the Competency Standards for Occupational Therapy Driver Assessors: An overview of the evidence for the inclusion of cognitive and perceptual assessments within fitness-to-drive evaluations. Australian Occupational Therapy Journal. 2017;64(4):328-39.

82. Gooden JR, Ponsford JL, Charlton JL, Ross PE, Marshall S, Gagnon S, et al. The development and initial validation of a new tool to measure self-awareness of driving ability after brain injury. Australian Occupational Therapy Journal. 2017;64(1):33-40.

83. Gooden JR, Ponsford JL, Charlton JL, Ross P, Marshall S, Gagnon S, et al. Self-regulation upon return to driving after traumatic brain injury. Neuropsychological Rehabilitation. 2016:1-15.

84. Wolfe PL, Lehockey KA. Neuropsychological Assessment of Driving Capacity. Archives of Clinical Neuropsychology. 2016;31(6):517-29.

85. Devlin A, McGillivray JA. Self-regulation of older drivers with cognitive impairment: A systematic review. Australasian Journal on Ageing. 2014;33(2):74-80.

86. Kay LG, Bundy AC, Clemson LM. Awareness of Driving Ability in Senior Drivers With Neurological Conditions. American Journal of Occupational Therapy. 2009;63(2):5.

87. DriveSafe DriveAware Research Summary. Pearson Australia Group Pty Ltd; 2015.

88. O'Donnell JM, Morgan MK, Manuguerra M. Functional outcomes and quality of life after microsurgical clipping of unruptured intracranial aneurysms: a prospective cohort study. Journal of neurosurgery. 2018 Feb:1-8.

89.   Batista JA, Giatti L, Barreto SM, Galery ARP, Passos VMdA. Reliability of Cognitive Tests of ELSA-Brasil, the Brazilian Longitudinal Study of Adult Health. Dementia & Neuropsychologia. 2013;7:367-73.

90.   Collie A, Maruff P, Darby DG, McStephen M. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. Journal of the International Neuropsychological Society. 2003;9(3):419-28.

91.   Prensky M. Digital natives, digital immigrants. Gifted. 2005(135):29-31.

92.   Wagner N, Hassanein K, Head M. Computer use by older adults: A multi-disciplinary review. Computers in Human Behavior. 2010;26(5):870-82.

93.   Jayroe TJ, Wolfram D. Internet searching, tablet technology and older adults. Proceedings of the American Society for Information Science and Technology. 2012;49(1):1-3.

94.   Lim FS, Wallace T, Luszcz MA, Reynolds KJ. Usability of Tablet Computers by People with Early-Stage Dementia. Gerontology. 2013;59(2):174-82.

95.   von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. The Lancet. 2007;370(9596):1453-7.

96.   Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Archives of Orofacial Science. 2017;12(1).

97.   Manuguerra M, Heller Gillian Z. Ordinal Regression Models for Continuous Scales. The International Journal of Biostatistics 2010.

98.   Kay LG, Bundy AC, Clemson LM. Validity, reliability and predictive accuracy of the Driving Awareness Questionnaire. Disability and Rehabilitation. 2009;31(13):1074-82.

99.   Wu Z, Ayton LN, Guymer RH, Luu CD. Intrasession Test–Retest Variability of Microperimetry in Age-Related Macular Degeneration. Investigative Ophthalmology & Visual Science. 2013;54(12):7378-85.

100.   Heeg GP, Ponsioen TL, Jansonius NM. Learning effect, normal range, and test–retest variability of Frequency Doubling Perimetry as a function of age, perimetric experience, and the presence or absence of glaucoma. Ophthalmic and Physiological Optics. 2003;23(6):535-40.

101.   Bentley SA, LeBlanc RP, Nicolela MT, Chauhan BC. Validity, Reliability, and Repeatability of the Useful Field of View Test in Persons with Normal Vision and Patients with Glaucoma. Investigative Ophthalmology & Visual Science. 2012;53(11):6763-9.

102.   Centofanti MMDP, Fogagnolo PMD, Oddone FMD, Orzalesi NMD, Vetrugno MMD, Manni GMD, et al. Learning Effect of Humphrey Matrix Frequency Doubling Technology Perimetry in Patients With Ocular Hypertension. Journal of Glaucoma. 2008;17(6):436-41.

103. Callesen J, Richter C, Kristensen C, Sunesen I, Næsby M, Dalgas U, et al. Test–retest agreement and reliability of the Six Spot Step Test in persons with multiple sclerosis. Multiple Sclerosis Journal. 2017;0(0):1352458517745725.

104. Casey AF, Wang X, Osterling K. Test-Retest Reliability of the 6-Minute Walk Test in Individuals With Down Syndrome. Archives of Physical Medicine and Rehabilitation. 2012;93(11):2068-74.

105. Rinne MB, Pasanen ME, Miilunpalo SI, Oja P. Test-Retest Reproducibility and Inter-Rater Reliability of a Motor Skill Test Battery for Adults. Int J Sports Med. 2001;22(03):192-200.

106. Santiago-Ramajo S, Katati MJ, Pérez-García M, Arjona-Moron V. Evaluating the Recovery of Cognitive Impairment in Subarachnoid Hemorrhage Taking Into Consideration the Practice Effects. Neurosurgery. 2010;67(6):1497-504.

107. Oberlander TJ, Olson BL, Weidauer L. Test-Retest Reliability of the King-Devick Test in an Adolescent Population. Journal of Athletic Training. 2017;52(5):439-45.

108. Sakzewski L, Lewis M, Ziviani J. Test–retest reproducibility of the Assessment of Motor and Process Skills for school-aged children with acquired brain injuries. Scandinavian Journal of Occupational Therapy. 2017;24(3):161-6.

109. Little CE, Emery C, Black A, Scott SH, Meeuwisse W, Nettel-Aguirre A, et al. Test–retest reliability of KINARM robot sensorimotor and cognitive assessment: in pediatric ice hockey players. Journal of NeuroEngineering and Rehabilitation. 2015;12(1):78.

110. Lundeberg MA, Fox PW, Punćochaŕ J. Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments. Journal of Educational Psychology. 1994;86(1):114-21.

111. Hammeke TA, Golden CJ, Purisch AD. A Standardized, Short, and Comprehensive Neuropsychological Test Battery Based on the Luria Neuropsychological Evaluation. International Journal of Neuroscience. 1978;8(3):135-41.

Appendices:

Human Research Ethics Committee Approval letter

Office of the Deputy Vice-Chancellor
(Research)

Research Office
Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 4459
http://www.research.mq.edu.au/
ABN 90 952 801 237

MACQUARIE University
SYDNEY·AUSTRALIA

14 February 2017

Dear Dr Davidson

Reference No: 5201600915

Title: *Functional Outcomes following Neurosurgical Treatment for Brain Tumours*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)).

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated May 2015) (the *National Statement*).

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement*, which is available at the following website:

http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research

2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.

3. All adverse events, including events which might affect the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.

4. Proposed changes to the protocol and associated documents must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC (Medical Sciences) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

The HREC (Medical Sciences) wishes you every success in your research.

Yours sincerely

Professor Tony Eyers
Chair, Macquarie University Human Research Ethics Committee (Medical Sciences)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

References